

09/806701

REC'D 25 NOV 1999

WIPO

PCT



Kongeriget Danmark

Patent application No.: PA 1998 01321

Date of filing: 15 Oct 1998

Applicant: Biolmage A/S
Mørkhøj Bygade 28
DK-2860 Søborg

This is to certify the correctness of the following information:

The attached photocopy is a true copy of the following document:

- The specification, claims, abstract and drawings as filed with the application on the filing date indicated above.

By assignment dated 08 Sept 1999 and filed 09 Sept 1999, the application has been assigned to Biolmage A/S



Patent- og
Varemærkestyrelsen
Erhvervsministeriet

TAASTRUP 16 Nov 1999

Lizzi Vester

Lizzi Vester
Head of Section

**PRIORITY
DOCUMENT**

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

15 OKT. 1998

A method for preventing or treating adverse conditions which may be reduced or abolished by modulating the activity of one or more phosphodiesterases having the ability to cleave cGMP.

SUMMARY OF THE INVENTION

This application describes a method to identify novel chemical entities that can maintain or increase cGMP levels within living cells. Such compounds will specifically modulate the effectiveness of the cyclic GMP-cleaving phosphodiesterases, for example PDE5, found in smooth muscle cells of vascular tissue. The preferred mode of action being sought is dislocation of specific isoforms of these cyclic nucleotide phosphodiesterases (hereafter referred to as PDE;s) from their anchoring sites within cells, thereby reducing their specific effectiveness, not their enzymatic capacity.

In its broadest aspect, the present application relates to a novel method for preventing or treating, in an animal in need thereof, an adverse condition which may be reduced or abolished by modulating the activity of one or more PDE;s having the ability to cleave cGMP. The method comprises modulation of the specific effectiveness of PDE or PDE;s by modulating their spatial distribution within cells of the animal.

The PDE is chosen from the group consisting of PDE 1, PDE 2, PDE 5, PDE 6, and PDE 9. More specifically, the method relates to mPDE5 or cPDE5 and splice variants thereof. The animal with the adverse condition may be a mammal and preferably a human.

In one embodiment of the invention modulation of the specific effectiveness of the PDE is a dislocation of the PDE from a native location within the cell.

In another embodiment of the invention modulation of the specific effectiveness of the PDE involves a disruption of its targeting to a native location within the cell.

In another embodiment of the invention modulation of the specific effectiveness of the PDE involves interference with the redistribution of the PDE, the redistribution being associated with an increase or a decrease of the specific effectiveness of the PDE.

The modulation of the specific effectiveness of the PDE may involve both an up-regulation or a down-regulation of the effectiveness of the PDE to perform its function within the cell.

This patent application is associated with the patent application "An improved method..." enclosed hereby as appendix A. Appendix A is considered part of the application.

BACKGROUND

Very many extracellular signals are transduced via intracellular systems employing the cyclic nucleotides cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP) as intermediaries, or second messengers. The processes mediated by cAMP and cGMP include control of smooth muscle tone, learning, vision, cellular differentiation, control of pro-inflammatory mediator production and action, apoptosis, lipogenesis, glycogenolysis and gluconeogenesis, circadian rhythms, cardiac function, and mood control through noradrenergic potentiation.

Cyclic nucleotides are generated by adenylate and guanylate cyclases (ACs and GCs, respectively) from ATP and GTP, signal to cAMP- and cGMP-dependent protein kinases (cAKs and cGKs, respectively) and directly to certain ion channels. cAMP and cGMP are removed by PDE;s. The required specificity of signals generated by these systems arises from diversity of type, tissue-specific expression and intracellular placement of the enzymes involved. For instance there are nine isoforms of ACs known plus additional splice variants, soluble and membrane located forms of GCs, multiple isoforms of the cAK and cGK kinases, and very many isoforms of PDEs of which over 30 have been identified (Perry and Higgs, 1998; Houslay and Milligan, 1997; Beavo, 1995). Additional specificity arises from targeting particular signalling enzymes to restricted locations within cells; this is the function of scaffold and anchoring proteins, such as the AKAP family, and not only may they place enzymes close to their substrates, but may also serve to recruit multiple enzymes into functional signalling units (Pawson and Scott, 1997).

Two groups of GCs are known in mammals, the soluble ones and those that are membrane located. GCs from both groups are central to systemic control of blood pressure. Soluble GCs are expressed in almost all cell types of the cardiovascular system including cardiomyocytes, vascular smooth muscle cells (VSMCs), endothelial cells and platelets (Drewett and Garbers, 1994). Soluble GCs contain a prosthetic heme group which binds NO (and CO) and leads to activation of the enzyme: the vasoactive properties of NO are mediated through the cGMP pathway in this way. The membrane located GCs act as receptors for various ligands (among them, natriuretic peptides and guanylin). cGMP-mediated functions of the natriuretic hormone receptors include vascular smooth muscle relaxation as well as regulation of blood volume (Benner *et al.*, 1990).

cGMP interacts with a number of different effector proteins:

- a) with certain ion channels e.g. in photoreceptors and olfactory cells, also in heart and kidney (Lincoln & Cornwell, 1993; Biel *et al.*, 1994; Light *et al.*, 1990);
- b) with cGMP-dependent protein kinases (cGKI and cGKII), of which "cytosolic" cGKI predominates in the cardiovascular system and has at

least 2 splice variants, α and β . cGKI α has 10-fold higher affinity for cGMP than the β variant. Both cGKI variants are found in vascular smooth muscle (Keilbach *et al.*, 1992, Hofmann *et al.*, 1992);

c) at high concentrations, with cAMP-dependent protein kinases (cAK), which being similar to the cGKs have a certain affinity for cGMP, just as the reverse is also true (Vaandrager & de Jonge, 1996). The functional significance of this potential cross-talk between pathways is not yet fully known, but may be connected with the anti-proliferative effects of cGMP (Lincoln *et al.*, 1994);

d) with cGMP-modulated PDEs: cGMP binds to a non-catalytic site of PDE2 and lowers its K_m for cAMP, lowering the baseline level of cAMP achievable by the enzyme. PDE3 catalysis of cAMP is effectively inhibited by cGMP (Pyne *et al.*, 1987), thus in cells where PDE3 predominates, increased cGMP leads to increased cAMP.

Smooth muscle contracts following Ca^{2+} -calmodulin activation of myosin light chain kinase (MLCK). cGKI relaxes smooth muscle by lowering free cytoplasmic Ca^{2+} levels, but the principal means by which this is accomplished varies considerably between types of smooth muscle, animal species, and the nature of the contractile stimulus being antagonised (Vaandrager & de Jonge, 1996). cGKI has been implicated in: inhibition of G-protein activation of phospholipase C β ; activation of Ca^{2+} -ATPase activity at plasma membrane and sarcoplasmic reticulum (SR); hyperpolarisation of membrane potential through activation of Ca^{2+} -activated K^+ channels; inhibition of voltage operated Ca^{2+} channels; stimulation of the $\text{Na}^+/\text{Ca}^{2+}$ exchanger; inhibition of SR IP_3 receptors. All of these actions require that the normally cytoplasmic cGKs must find membrane located targets, and specific anchor proteins may be involved. cGKI is already known to be targeted to specific anchor proteins of the cytoskeleton (MacMillan-Crow & Lincoln, 1994), and the discovery of further interactions is likely.

Inactivation of cAMP/cGMP occurs by hydrolysis of the 3'-ester bond, catalysed by the PDEs. The major cGMP-degrading PDEs are types 1,2,5, 6 and 9 but here we focus on PDE5, since this is the principal cGMP-specific PDE found in airway and vascular smooth muscle, and it is one of the better documented families of cGMP-specific PDEs. Little is known yet concerning the role of the newly discovered PDE9 isoforms (Soderling *et al.*, 1998; Fisher *et al.*, 1998), and the situation is similar for PDE2s, since good inhibitors are as yet unknown for these (Perry and Higgs, 1998). PDE5 is activated by cAK and (10-fold faster) by cGK (Thomas *et al.*, 1990). Phosphorylation of PDE5 is enhanced in the presence of cGMP, and apparently increases the enzyme's

V_{\max} by 10-fold (Burns *et al.*, 1992). Coupled with PDE3, these interactions form a feedback system to limit cGMP signaling: increased cGMP will increase cAMP through inhibition of PDE3, high cAMP will activate cAK which, in the presence of elevated cGMP will activate PDE5 and therefore stimulate cGMP breakdown. cAMP levels return to baseline as cGMP falls, by re-activation of PDE3. Recent evidence (Pyne *et al.*, 1996; Lochhead *et al.*, 1997) suggests that PDE5 may have additional protein components associated with it analogous to the gamma subunits of PDE6. The PDE6 γ subunits serve to link activation of the G-protein transducin to activation of the PDE. They are subsequently involved in turning off the signal by helping to activate the transducin GTPase. In the case of PDE5, these associated proteins (14 to 18 kDa) may serve to block activation of the enzyme by cGK and cAK, and the blocking ability of these polypeptides appears to be controlled by a G-protein regulated kinase (Pyne *et al.*, 1996).

The PDEs are key components of the cyclic nucleotide signalling systems, allowing local concentration differences of the cyclic nucleotide messengers to be established, between adjacent tissues, between adjacent cells, even within a single cell between different volumes of cytoplasm. The ability to generate such heterogeneity in the distribution of concentrations of a commonly shared signalling molecule, such as cAMP, is at the heart of directed signalling processes. To be of therapeutic value, cyclic nucleotide control has to be achieved with defined cellular selectivity (Perry and Higgs, 1998). It is the therapeutic opportunities offered by certain of the PDEs that concerns this application.

Ten families of PDEs have been identified, designated simply PDE1 to PDE10. Within each family there are two or more related but distinct gene products (A, B, C, etc.) and for each of these alternative mRNA processing gives rise to multiple splice variants, identified by an additional arabic numeral in accordance with the most recent nomenclature recommendation (Molecular Pharmacology 46:399-405, 1994). All PDE gene products identified so far have two functional domains per molecule, one catalytic, and one regulatory. The catalytic domain lies towards the carboxylic acid terminus of each PDE protein and has the greatest homology between the PDE families, being >75% homologous at the amino acid level (Perry and Higgs, 1998). Nevertheless, each of the more than 30 PDEs known have individually distinct substrate specificities, kinetic characteristics, regulatory properties and cellular and subcellular distributions (Houslay and Milligan, 1997).

PDEs 4, 7 and 8 are highly specific for cAMP. PDEs 5, 6, and 9 are selective for cGMP. PDE3s bind cAMP and cGMP with similar affinity, but hydrolyse cAMP most efficiently, cGMP rather poorly. PDE3s are therefore negatively regulated in their

cAMP hydrolysing ability by cGMP. PDEs 1 and 2 hydrolyse both cAMP and cGMP, but with PDE1 the relative efficiencies vary with isoenzyme subtype (Perry and Higgs, 1998).

The amino terminal ends of PDEs consist of the regulatory domains, which are very different both between families and between variants within families. This region contains variously: a binding domain for Ca^{2+} -calmodulin (CaM) in PDE1; non-catalytic cGMP-binding sites in PDEs 2, 5 and 6; a binding domain for the signalling G-protein transducin in PDE6. The amino terminal region also contains membrane-targeting sequences in several PDE3s and PDE4s, as well as protein kinase phosphorylation sites in PDEs 1, 3, 4 and 5. These phosphorylation sites are likely to be important in regulation of catalytic activity and/or subcellular location (Perry and Higgs, 1998).

This application concerns the particular therapeutic benefits arising from selective inhibition of PDE isoforms and variants which hydrolyse cGMP.

Blood pressure elevation to a degree that requires medical treatment is often encountered in up to 15% of an adult population. In only 10-15% of these, a definite cause for the hypertension can be found and in the rest, the "essential hypertension" has to be treated without a hope for cure of the underlying disease. Long-standing elevation of blood pressure, even quite moderate, damages vessels in the heart, kidneys and brain and dramatically increases the risk for coronary heart disease, renal failure and stroke. It has been shown that effective pharmacologic treatment of hypertension substantially reduces morbidity and mortality from these conditions. The finding that endothelial cells produce a local vascular relaxation factor, identified as nitric oxide (NO), that activates guanylyl cyclase and increases cGMP that in turn leads to reduction in vascular smooth muscle cell tone, has opened new possibilities for blood pressure regulation / vasorelaxation based on modulation of the cellular levels of cGMP. A number of the components in the cGMP system displays tissue specific distribution (Vaandrager & de Jonge, 1996; Pyne *et al.*, 1996). This increases the likelihood for improved pharmacological specificity and fewer side-effects when using these as targets for antihypertensive treatment instead of the traditional ones. It is the cGMP-dependent protein kinase (PKG) (Vaandrager & de Jonge, 1996) that is thought to mediate the intracellular effects of cGMP. The cGMP -dependent and -specific phosphodiesterases can serve as connectors to the cAMP system and terminators of cGMP effects (Pyne *et al.*, 1996). PDE5 has attracted attention since it is selective for degradation of cGMP versus cAMP. Isoform-specific inhibitors for PDE5 are being developed by several companies and one compound from Pfizer, Sildenafil, has proven selectivity for PDE5 and is currently being marketed as treatment against impotence (Viagra), originally a side-effect resulting from vasorelaxation in the corpus cavernosum. However the

screening procedures currently used search only for direct enzymatic inhibitors of PDE and the compounds found are often not selective, inhibiting for instance both PDE 1 and 5 (e.g. Zaprinast (M&B 22948 RPR), Sch 59498 and Sch 51866). By the methods described herein and within appendix A, new chemical entities can be found which primarily will be specific modulators of PDE action, not inhibitors of the enzymatic action *per se*. Preferred compounds will inhibit the site-specific anchoring of PDEs which hydrolyse cGMP, and thereby reduce their effectiveness in controlling local concentrations cGMP within living cells.

The therapeutic potential of selective modulators of cGMP-related PDE action is not restricted to relaxation of smooth muscle cells but also encompasses other effects ascribed to PKG, such as inhibition of platelet activation (Chiu *et al.*, 1997; Vemulapalli *et al.*, 1996), inhibition of endothelial permeability increases in response to vasoactive substances (Raeburn & Karlsson, 1993), inhibition of the differentiation of osteoclasts (Holliday *et al.*, 1997) and light-induced resetting of circadian rhythms (Mathur *et al.*, 1996; Liu *et al.*, 1997).

The search for chemical inhibitors of the catalytic activity of specific PDEs is currently one of the most intensive areas of pharmaceutical research, particularly so for PDEs 4 and 5. Much progress has been made in this area, with several compounds known to have selective activity for particular families of PDEs (reviewed in Perry and Higgs, 1998; Hughes *et al.*, 1997; Teixeira *et al.*, 1997). However, there has not yet been found a class of compounds able to select between isoenzymes within the same family, which is where the greatest opportunities lie. Without isoform specificity, certain difficulties can be expected with the use of enzymic inhibitors of PDEs. Some of these difficulties are outlined below.

In general, the effects a known inhibitor of the catalytic activity of a particular class of PDEs may have on cyclic nucleotide levels often varies between different cell types. The reasons for this are several, but include: differences in the basal level of cyclase activity in distinct cell types, crosstalk between cAMP and cGMP systems, and differences in local concentrations of substrate within a cell which influences the degree of inhibition that can be attained by a simple competitive enzyme inhibitor (Perry and Higgs, 1998). First, PDE inhibition is only useful if it produces the appropriate change in the activity of the dependent effectors, for instance activation of cAK when the concentration of cAMP can be increased above a threshold level. The rate of change in concentration depends in part on the activity of the cyclases which generate the cyclic nucleotides, and that basal level of activity differs from isoform to isoform, and therefore from cell type to cell type. In adipocytes, for example, AC activity is high and cAMP levels are kept at baseline only by a correspondingly high PDE activity. Hepatocytes on the other hand

have a rather low AC activity. If both cell types share PDEs of the same family, and are treated with a chemical inhibitor targeting that family, there will be a rapid increase in cAMP within adipocytes and activation of their cAKs, but no activation in hepatocytes, unless the AC is also stimulated.

Second, general inhibition of a particular isoform of PDE can have certain unavoidable consequences on other cyclic nucleotide pathways since cAMP and cGMP systems are often closely interlinked. Much of this crosstalk arises from PDE regulation by cyclic nucleotides. When cGMP increases in platelets (e.g. following nitric oxide stimulation of soluble GC, or PDE5 inhibition) it inhibits PDE3 and causes a concomitant rise in cAMP (Ashida and Sakuma, 1992). In adrenal glomerulosa cells, atrial natriuretic factor elevates cGMP but inhibits cAMP-stimulated aldosterone synthesis via cGMP-stimulation of PDE2 (MacFarland *et al.*, 1991).

Third, the expected effects of PDE inhibition may be modified by differences in local concentrations of substrates, the reason being that most chemical inhibitors of PDE action are competitive with substrate, so their therapeutic profile is dependent on both the Michaelis-Menton equilibrium constant (K_M) and the substrate concentration in which they are operating (Perry and Higgs, 1998). Most effective inhibition will always occur at lowest substrate levels, but as a corollary, a locally increased substrate level will reduce the inhibition attained. In combination with subtle differences in isoform K_M values for an inhibitor, the desired spatial modulation of cyclic nucleotide levels within a cell could be difficult to obtain by simple competitive inhibition of catalytic activity.

The radically different methods of interference with PDE action as proposed below in this application should avoid many of the problems outlined above, principally because interference will be family and isoform specific and targeted not against catalytic activity of the PDEs, but their spatial organisation within the cell.

Targeting of signalling enzymes is a recognised mechanism by which sensitivity, specificity, precision and control may be introduced into intracellular signalling pathways (Pawson and Scott, 1997; Faux and Scott, 1996). The importance and occurrence of targeting as a phenomenon are described and discussed in appendix A. Of central importance to this application is the modulation of the effectiveness of signalling PDEs through interference with their intracellular targeting. As already described, the many PDEs known share much structural homology, and this is especially true within the catalytic regions found towards the carboxylic acid terminals of the proteins. At the amino terminals much more heterogeneity is found, between families of PDEs, between isoforms within families, and between splice variants derived from individual gene isoforms (Houslay and Milligan, 1997). Much of this heterogeneity appears to be associated with differences in targeting behaviour, at least in PDE4 isoforms and

variants (Scotland *et al.*, 1998, Bolger *et al.*, 1997), and by extension should apply to other PDEs as well since they are in overall character similar protein molecules with similar roles in cellular signalling.

Evidence suggests that the amino terminal regions of PDEs can serve to target isoforms to specific intracellular sites (Shakur *et al.*, 1995; McPhee *et al.*, 1995; Bolger *et al.*, 1996; Pooley *et al.*, 1997) and that they can regulate the functioning of the catalytic unit either through interaction with binding proteins (Shakur *et al.*, 1995; O'Connell *et al.*, 1996; Pyne *et al.*, 1996) or through phosphorylation (Sette and Conti, 1996). Targeting appears to occur through protein-protein interactions with membrane- or cytoskeletally-located proteins (Houslay, Sullivan and Bolger, 1998), and of these the membrane associated proteins include both integral and peripherally adherent species. Such interactions have been probed at a gross level through the use of nonionic detergents and elevated ionic strength (Scotland *et al.*, 1998).

So far only two PDE5 genes are known and two enzyme variants have been reported. In parallel with other PDE isoforms more splicing variants are to be expected from each gene. The enzyme is a homodimer, each subunit being 93 kDa. The structural organisation of the dimer is very similar to that of the cGKs.

PDE5s exist in two distinct forms: one membrane-bound (mPDE5) and one cytosolic (cPDE5) (Pyne *et al.*, 1996). The mPDE5 is activated by PKA and is inhibited by a G-protein dependent mechanism. It is assumed that cPDE5 is part of a "signalling cassette" with NO-regulated guanylate cyclase and PDE3. The latter construction will lead to very short-lived messages whereas the former allows for generation of prolonged cGMP signals

Targeting or anchoring of PDE5s is likely to have its greatest effect through compartmentalisation of cGMP signalling within cells. Associated with the PDE5s will be specific GCs together with specific isoforms of the effector cGK, or cGMP-operated ion channels. cGKs may be attached to specific G-kinase anchoring proteins. Specific subcellular distributions of these components will allow for spatial and temporal gradients of cGMP to be established within cellular compartments. Targeted PDE5 species might serve to control threshold levels of cGMP in the environs of specific cGK molecules, perhaps protecting certain protein complexes from cGK-mediated phosphorylation or manipulating the activity levels of GCs that are necessary before cGK activation may occur.

Competitive chemical inhibitors are known which can selectively inhibit PDE5s. Relatively few isoforms of PDE5 are known to date. PDE5 is found rather specifically in vascular and airway smooth muscle. That sildenafil, with its 5 nM IC₅₀ for PDE5,

affects only a subset of vascular smooth muscle is puzzling, but strongly suggests that either multiple PDE5 isoforms or states exist in different vascular smooth muscle, presumably with different sensitivities to sildenafil, or more likely, other cGMP-hydrolysing PDEs are important in different vascular smooth muscles.

As to other potentially important cGMP-hydrolysing PDE targets, many are doubtless yet to be discovered. PDE9s have only been known since the end of 1997. These have a rather general distribution (kidney, brain, lung), have a very high affinity for cGMP (about 70 nM) and are inhibitable by the PDE1/5 inhibitor SCH51866 (1.55 μ M), but "not by sildenafil" (7 μ M, Soderling *et al.*, 1998). Their physiological roles and regulation have not been defined (Soderling *et al.*, 1998; Fisher *et al.*, 1998), but the best suggestions are that they may be involved in keeping cGMP at very low levels when activated, and may, in kidney, be involved in termination of ANP signalling, and therefore inhibition may help potentiate natriuresis without causing deleterious drops in blood pressure (Soderling *et al.*, 1998).

It is clear that PDEs possess heterogeneity in their amino terminal regions, and the approach outlined in this application exploits those differences between isoforms and splice variants to produce what should be confined and defined therapeutic effects. Furthermore, in many cases it may be expected that dislocation of an active enzyme from a targeted site of action will have little effect on average cellular concentrations of cGMP, although significant increases may occur at the now PDE-free site of action. This may have significance where an acute short-term process is the therapeutic target, but an integrative gene-regulation effect may occur upon general, non-specific PDE inhibition and overall cGMP increase in the cell.

DETAILED DISCLOSURE

In the present specification and claims, the term "influence" covers any influence to which the cellular response comprises a redistribution. Thus, e.g., heating, cooling, high pressure, low pressure, humidifying, or drying are influences on the cellular response on which the resulting redistribution can be quantified, but perhaps the most important influence is the influence of contacting or incubating the cell or cells with a substance which is known or suspected to cause a redistribution. In another embodiment of the invention the influence could be substances from a compound drug library.

In the present context, the term "green fluorescent protein" (GFP) is intended to indicate a protein which, when expressed by a cell, emits fluorescence upon exposure to light of the correct excitation wavelength (cf. Chalfie, M. *et al.* (1994) Science 263, 802-805).

In the following, GFP in which one or more amino acids have been substituted, inserted or deleted is also termed "modified GFP". "GFP" as used herein includes wild-type GFP derived from the jelly fish *Aequorea victoria* and modifications of GFP, such as the blue fluorescent variant of GFP disclosed by Heim et al. (Heim, R. *et al.* (1994). Proc.Natl.Acad.Sci. 91:26, pp 12501-12504), and other modifications that change the spectral properties of the GFP fluorescence, or modifications that exhibit increased fluorescence when expressed in cells at a temperature above about 30°C described in PCT/DK96/00051, published as WO 97/11094 on 27 March 1997 and hereby incorporated by reference, and which comprises a fluorescent protein derived from *Aequorea* Green Fluorescent Protein or any functional analogue thereof, wherein the amino acid in position 1 upstream from the chromophore has been mutated to provide an increase of fluorescence intensity when the fluorescent protein of the invention is expressed in cells. Preferred GFP variants are F64L-GFP, F64L-Y66H-GFP and F64L-S65T-GFP. An especially preferred variant of GFP for use in all the aspects of this invention is EGFP (DNA encoding EGFP which is a F64L-S65T variant with codons optimized for expression in mammalian cells is available from Clontech, Palo Alto, plasmids containing the EGFP DNA sequence, cf. GenBank Acc. Nos. U55762, U55763).

The terms "intracellular signalling pathway" and "signal transduction pathway" are intended to indicate the coordinated intracellular processes whereby a living cell transduces an external or internal signal into cellular responses. Said signal transduction will involve an enzymatic reaction said enzymes include but are not limited to protein kinases, GTPases, ATPases, protein phosphatases, phospholipases and cyclic nucleotide phosphodiesterases. The cellular responses include but are not limited to gene transcription, secretion, proliferation, mechanical activity, metabolic activity, cell death.

The term "second messenger" is used to indicate a low molecular weight component involved in the early events of intracellular signal transduction pathways.

The term "luminophore" is used to indicate a chemical substance which has the property of emitting light either inherently or upon stimulation with chemical or physical means. This includes but is not limited to fluorescence, bioluminescence, phosphorescence, chemiluminescence.

The term "mechanically intact living cell" is used to indicate a cell which is considered living according to standard criteria for that particular type of cell such as maintenance of normal membrane potential, energy metabolism, proliferative capability, and has not experienced any physically invasive treatment designed to introduce external substances into the cell such as microinjection.

In the present context, the term "permeabilised living cell" is used to indicate cells where a pore forming agent such as Streptolysin O or *Staphylococcus Aureus* α -toxin has been applied and thereby incorporated into the plasma membrane in the cells. This creates proteinaceous pores with a defined pore size in the plasma membranes of the exposed cells. Pores could also be made by electroporation, i.e. exposing the cells to high voltage discharges, a procedure that creates small holes in the plasma membrane by coagulating integral membrane proteins. Treatment with a mild detergent such as saponin may accomplish the same thing. Common to all these treatments is that pores are formed only in the plasma membrane without affecting the integrity of cytoplasmic structural elements and organelles. The term living in this context means that the permeabilised cell or cells bathed in a solution mimicking the intracellular milieu still have functional organelles, such as actively respiring mitochondria and endoplasmatic reticulum that can take up and release calcium ions, and functional structural elements. In one embodiment this method is applied so that substances that normally can not traverse the plasma membrane, but most likely exert their influence intracellularly, can be introduced and their influence studied. In another embodiment this method is used to record the response to an influence from many cells simultaneously.

In the present context, the term "permeabilisation" is intended to indicate the selective disruption of the plasma membrane barrier so that soluble substances freely mobile in the cytosol may be lost from the interior of the cells. The permeabilisation can be achieved as described above under "permeabilised living cells" or by using other chemical detergents such as Triton X-100 or digitonin in carefully titrated amounts.

The term "physiologically relevant", when applied to an experimentally determined redistribution of an intracellular component, as measured by a change in the luminescence properties or distribution, is used to indicate that said redistribution can be explained in terms of the underlying biological phenomenon which gives rise to the redistribution.

The terms "image processing" and "image analysis" are used to describe a large family of digital data analysis techniques or combination of such techniques which reduce ordered arrays of numbers (images) to quantitative information describing those ordered arrays of numbers. When said ordered arrays of numbers represent measured values from a physical process, the quantitative information derived is therefore a measure of the physical process.

The term "mammalian cell" is intended to indicate any living cell of mammalian origin. The cell may be an established cell line, many of which are available from The American Type Culture Collection (ATCC, Virginia, USA) or a primary cell with a limited life span derived from a mammalian tissue, including tissues derived from a transgenic

animal, or a newly established immortal cell line derived from a mammalian tissue including transgenic tissues, or a hybrid cell or cell line derived by fusing different celltypes of mammalian origin e.g. hybridoma cell lines. The cells may optionally express one or more non-native gene products, e.g. receptors, enzymes, enzyme substrates, prior to or in addition to the fluorescent probe. Preferred cell lines include but are not limited to those of fibroblast origin, e.g. BHK, CHO, BALB, or of endothelial origin, e.g. HUVEC, BAE (bovine artery endothelial), CPAE (cow pulmonary artery endothelial), HLMVEC (human lung microvascular endothelial cells) or of pancreatic origin, e.g. RIN, INS-1, MIN6, bTC3, aTC6, bTC6, HIT, or of hematopoietic origin, e.g. primary isolated human monocytes, macrophages, neutrophils, basophils, eosinophils and lymphocyte populations, AML-193, HL-60, RBL-1, adipocyte origin, e.g. 3T3-L1, neuronal/neuroendocrine origin, e.g. AtT20, PC12, GH3, muscle origin, e.g. SKMC, A10, C2C12, renal origin, e.g. HEK 293, LLC-PK1.

The term "hybrid polypeptide" is intended to indicate a polypeptide which is a fusion of at least a portion of each of two proteins, in this case at least a portion of the green fluorescent protein, and at least a portion of a catalytic and/or regulatory domain of a protein kinase. Furthermore a hybrid polypeptide is intended to indicate a fusion polypeptide comprising a GFP or at least a portion of the green fluorescent protein that contains a functional fluorophore, and at least a portion of a biologically active polypeptide as defined herein provided that said fusion is not the Glucocorticoid Receptor-GFP disclosed by Carey, KL et al. and Guiliano, KA et al., respectively. Thus, GFP may be N- or C-terminally tagged to a biologically active polypeptide, optionally via a linker portion or linker peptide consisting of a sequence of one or more amino acids. The hybrid polypeptide or fusion polypeptide may act as a fluorescent probe in mechanically intact or permeabilised living cells carrying a DNA sequence encoding the hybrid polypeptide under conditions permitting expression of said hybrid polypeptide. The term hybrid polypeptide or fusion polypeptide is intended also to include the term "fluorescent probe", where the latter is used to indicate a fluorescent fusion polypeptide comprising a GFP or any functional part thereof which is N- or C-terminally fused to a biologically active polypeptide as defined herein, optionally via a peptide linker consisting of one or more amino acid residues, where the size of the linker peptide in itself is not critical as long as the desired functionality of the fluorescent probe is maintained. A fluorescent probe according to the invention is expressed in a cell and basically mimics the physiological behaviour of the biologically active polypeptide moiety of the fusion polypeptide.

The term "kinase" is intended to indicate an enzyme that is capable of phosphorylating a cellular component.

The term "protein kinase" is intended to indicate an enzyme that is capable of phosphorylating serine and/or threonine and/or tyrosine in peptides and/or proteins.

The term "phosphatase" is intended to indicate an enzyme that is capable of dephosphorylating phosphoserine and/or phosphothreonine and/or phosphotyrosine in peptides and/or proteins.

The term "cyclic nucleotide phosphodiesterase" is intended to indicate an enzyme that is capable of inactivating the second messengers cAMP and cGMP by hydrolysis of their 3'-ester bond.

In the present context, the term "biologically active polypeptide" is intended to indicate a polypeptide affecting intracellular processes upon activation, such as an enzyme which is active in intracellular processes or a portion thereof comprising a desired amino acid sequence which has a biological function or exerts a biological effect in a cellular system. In the polypeptide one or several amino acids may have been deleted, inserted and/or replaced to alter its biological function, e.g. by rendering a catalytic site inactive or by disrupting the targeting sequence. In another embodiment, one or several amino acids may have been deleted, inserted and/or replaced without altering the biological function of the polypeptide, that is, it remains biologically equivalent. Preferably, the biologically active polypeptide is selected from the group consisting of proteins taking part in an intracellular signalling pathway, such as enzymes involved in the intracellular phosphorylation and dephosphorylation processes including kinases, protein kinases and phosphorylases as defined herein, but also proteins making up the cytoskeleton play important roles in intracellular signal transduction and are therefore included in the meaning of "biologically active polypeptide" herein. More preferably, the biologically active polypeptide is a protein which according to its state as activated or non-activated changes localisation within the cell, preferably as an intermediary component in a signal transduction pathway. Included in this preferred group of biologically active polypeptides are cAMP dependent protein kinase A and cyclic nucleotide phosphodiesterases.

The term "a substance" is intended to indicate any sample which has a biological function or exerts a biological effect in a cellular system. The sample may be a sample of a biological material such as a sample of a body fluid including blood, plasma, saliva, milk, urine, or a microbial or plant extract, an environmental sample containing pollutants including heavy metals or toxins, or it may be a sample containing a compound or mixture of compounds prepared by organic synthesis or genetic techniques.

The phrase "any change in fluorescence" means any change in absorption properties, such as wavelength and intensity, or any change in spectral properties of the emitted

light, such as a change of wavelength, fluorescence lifetime, intensity or polarisation, or any change in the intracellular localisation of the fluorophore. It may thus be localised to a specific cellular component (e.g. organelle, membrane, cytoskeleton, molecular structure) or it may be evenly distributed throughout the cell or parts of the cell.

The term "organism" as used herein indicates any unicellular or multicellular organism preferably originating from the animal kingdom including protozoans, but also organisms that are members of the plant kingdoms, such as algae, fungi, bryophytes, and vascular plants are included in this definition.

The term "nucleic acid" is intended to indicate any type of poly- or oligonucleic acid sequence, such as a DNA sequence, a cDNA sequence, or an RNA sequence.

The term "biologically equivalent" as it relates to proteins is intended to mean that a first protein is equivalent to a second protein if the cellular functions of the two proteins may substitute for each other, e.g. if the two proteins are closely related isoforms encoded by different genes, if they are splicing variants, or allelic variants derived from the same gene, if they perform identical cellular functions in different cell types, or in different species. The term "biologically equivalent" as it relates to DNA is intended to mean that a first DNA sequence encoding a polypeptide is equivalent to a second DNA sequence encoding a polypeptide if the functional proteins encoded by the two genes are biologically equivalent.

The term "fixed cells" is used to mean cells treated with a cytological fixative such as glutaraldehyde or formaldehyde, treatments which serve to chemically cross-link and stabilize soluble and insoluble proteins within the structure of the cell. Once in this state, such proteins cannot be lost from the structure of the now-dead cell.

In the present context a "quantitative fluorescence redistribution assay" is intended to indicate an assay whereby it is possible to observe and quantify the subcellular localisation and possible redistribution of an biologically active polypeptide, or part thereof, genetically or chemically tagged with a luminophore inside an intact living cell or cells or permeabilised living cells. The subcellular location and redistribution may be monitored using fluorescence microscopy or fluorescence imaging microscopy but is preferably monitored using a fluorescence imaging plate reader or a fluorescence plate reader for improved throughput. A more thorough description is given in Appendix A.

In the present context a "mortal cell line" is used to indicate animal cells that may grow in vitro, given the right conditions, but that have a definite life span of a number of cell divisions or days, week or months beyond which it is not at present possible to keep them alive.

In the present context an "immortalised cell line" is used to indicate cells of animal origin where the normal limitations for cell life and number of cell divisions do not

apply. Essentially, such cells can live, grow and divide for an unlimited or very long (years to decades) time.

The term "targeting sequence" is used to indicate the amino-acid sequence of a biologically active polypeptide that contains the actual structure or structures necessary for association of the biologically active polypeptide with its native intracellular binding sites. The term "targeting sequence" is also used to indicate the amino-acid sequence of a protein that contains the actual structure or structures necessary for association of a biologically active polypeptide with the protein.

The term "targeting" is used to indicate the process whereby a spatially distributed protein is directed to the intracellular sites and maintained at the intracellular sites to which it is normally anchored or associated. These anchoring sites are normally assumed to be the intracellular sites where the protein has its optimal function for the cell.

The term "dislocate" and derivatives thereof is used to indicate the process whereby an intracellularly spatially distributed protein is forced to detach from its normal anchoring or association structures in the cells due to intercalation of another, preferably smaller, compound at the site of anchoring or association. This usually means that the optimal function of the protein within the cell is lost or reduced and that a larger portion of the protein molecules are freely mobile within the cytoplasm.

In the present context a "screening assay" is intended to mean any measurement protocol, including materials, cells, instruments, chemicals, reagents, detection units, calibration and quantification procedures used to measure a response from mechanically intact or permeabilised living cells relevant to influences on an intracellular pathway.

In the present context a "primary screening assay" is used to indicate the first screening assay in a discovery project that is used to select and sort all compounds available to the project according to the quantified effect of the compounds in the assay.

In the present context a "counterscreen" is intended to mean a screening assay that is relevant to a phenomenon that is undesirable seen from the point of view of the discovery project.

In the present context a "discovery project" is intended to mean the process whereby general or specific ideas about ways of how to modulate an intracellular signalling pathway are exploited in order to find new chemical compounds that can be used to modulate the intracellular signalling pathway and thereby treat, reduce or abolish symptoms associated with a condition or a disease that is lethal, degenerative, performance-reducing or just uncomfortable to an animal, preferably a human being. The aim of the discovery project is to produce drug candidates that can be tested as potential drugs in an animal, preferably in human beings. The term "discovery project" also encompasses the actual group of individuals, screening assays, tests, machinery, cells, animals and compounds involved in different aspects of the project.

The term "tagging" is used to indicate the process whereby a luminophore is genetically or chemically attached to the protein, or part of the protein, of interest to the discovery project.

The term "primary hit" is used to indicate compounds identified in the primary screening assay as having at least the minimum level of desired effect that has been specified in the discovery project.

The term "primary lead compound" is used to indicate a primary hit that has at least the minimal level of desired potency and specificity predetermined by the discovery project.

The term "dose-response relationship" is in the present context intended to mean a clear correlation between the quantified response of cells in a screening assay to application of an influence, such as a compound, and the concentration of the applied influence. The response to the influence may be both an up-regulation and a down-regulation of the quantitated parameter used in the screening assay.

In the present context, the term "potency" is intended to mean the ability of an influence to affect the process under study. The process under study may be, for example a screening assay or a specific physiological or patophysiological response in an animal.

In the present context, the term "selectivity" is intended to mean the difference in potency on the desired process, such as a screening assay, and an undesired process, such as a counterscreen, with the view of the discovery project. An influence or a compound is said to display selectivity if the potency for the desired process is higher than for the undesired process.

In the present context, the term "structure-activity relationship" or "SAR" is intended to mean the situation where a direct relationship exists between a compound and modifications made to the compound and the activity of the compound and the modifications made to the compound in one or more screening assays. The process of building a SAR may be used to direct the chemical construction of new compounds with higher potency and selectivity than the original compound.

The term "drug candidate lead" is used to indicate compounds that may be pursued by a discovery project as potential candidates for the final outcome of the project.

In the present context, the term "efficacy" is intended to mean the ability of a compound to affect the process or condition under study. It is closely related to the term "potency" but is in the present context used when relating to effects of a compound on more complex screening assays than the primary screening assay or counterscreens and when relating to effects of a compound in animals.

In the present context, the term "toxicity" is intended to mean that a compound in some way is toxic to cells, tissues or animals. The toxicity means that the cells, tissues or animals will in some way be harmed if the compound is applied at a sufficient concentration. The effects may ultimately lead to cell, tissue or animal death or a limited life compared to the normal condition.

In the present context, the term "physiology" is intended to mean the normal function of biological and biochemical processes inside cells, between cells and in the whole organism or animal.

In the present context, the term "pathophysiology" is intended to mean deviations from the normal function of biological and biochemical processes inside cells, between cells and in the whole organism or animal that may be part of a condition or disease.

In the present context, the term "patogenesis" is intended to mean the process, be it genetical, biological, biochemical, chemical or environmental, that ultimately may explain, at least in part, the apparent patophysiology associated with a condition or disease in an animal.

In the present context, the term "fractionated cells" is intended to mean the outcome of a simple division of initially mechanically intact living cells into two fractions, particulate (the components that can be sedimented by centrifugation at more than 10 000xg and not more than 100 000xg for 10 minutes) and soluble fraction (the soluble components and small membrane fragments that do not sediment), after subjecting the cells to plasma membrane disruption either mechanically with some form of homogeniser or sonicator or osmotically (hypoosmotic shock) or through some kind of permeabilisation of the plasma membrane with detergents, toxins or electroporation.

The term "parenteral route of administration" is used to indicate the administration of a drug or compound in solution to an animal, such as a mammal or a human, by injection or infusion of the drug or compound into the bloodstream of the animal via an injection needle inserted into one of the animal's blood vessels, preferably a vein.

The term "oral route of administration" is used to indicate the administration of a drug or compound in solution or as a solid to an animal, such as a mammal or a human, by placing the drug or compound in the mouth of the animal so that the animal itself can swallow the drug or compound or have it delivered to the stomach or intestine by intubation. When the drug or compound enters the stomach and intestine it will be taken up over the mucosa into the bloodstream and administered via the blood stream to the tissues and organs where it is to exert its effect, or it will be acting locally in the stomach and intestine.

The term "pulmonary route of administration" is used to indicate the administration of a drug or compound as an aerosol with either solid or liquid particles to an animal, such as a mammal or a human, by placing the drug or compound container close to or in contact with the mouth and/or nose of the animal so that the animal itself can inhale the drug or compound aerosol. When the drug or compound enters the peripheral bronchioles and alveoli it will be taken up over the alveolar membrane, either into the bloodstream and administered via the blood stream to the tissues and organs where it is to exert its effect or it will act locally in the lungs on lung, vessel and muscle cells as well as any other cell type present there.

The term "cutaneous route of administration" is used to indicate the administration of a drug or compound in solution or as a solid to an animal, such as a mammal or a human, by placing the drug or compound on the skin of the animal. The drug can then enter the blood vessels under the skin as it is permeating the skin and thereby be taken up into the bloodstream and administered via the blood stream to the tissues and organs where it is to exert its effect. It may also exert an effect locally on the site of application on the skin.

The term "rectal route of administration" is used to indicate the administration of a drug or compound in solution or as a solid to an animal, such as a mammal or a human, by placing the drug or compound in the rectal cavity of the animal. When the drug or compound enters the rectum and parts of the large intestine it will be taken up over the mucosa into the bloodstream and administered via the blood stream to the tissues and organs where it is to exert its effect, or it will act locally in the rectum and parts of the large intestine.

Very many phosphodiesterases (PDE;s) are known. They are grouped in families according to functional criteria. Within each family there may be several members - isoforms- encoded by different genes. Each isoform may give rise to several splice variants. This hierarchy is evidenced at the sequence level: isoforms are more similar to each other than to members of other families; splice variants are more similar to each other than to other PDE;s. Each specific PDE thus contains sequences that are unique to itself, as well as sequences that are shared between isoforms and/or families.

When setting up a program to identify pharmacological agents that affect the intracellular distribution of a target PDE, it is first necessary to choose the target among the many PDE;s known. This may be done according to various criteria. A preferred embodiment is that the target PDE be present in the tissue or cell type(s) where the pharmacological agent is to exert its effect. In another preferred embodiment it is desirable that the target not be present in tissues or cell types where no pharmacological effects are desired.

Establishing the expression patterns of PDE;s in relation to tissues and cell types is best done using the methods of detection of mRNA, e.g. Northern analysis, which is a well established procedure. Briefly, mRNA isolated from a given source is probed with a labelled nucleotide, whose sequence is complementary to the mRNA or a region in a mRNA of interest. The assay allows the investigator to determine the stringency of the probing, i.e. to correlate the resulting signal(s) with sequence similarities.

As a first step, the nucleotide sequences of PDE;s are compiled and inspected to identify regions that are unique to specific PDE;s as well as regions that are shared among several, many, or all PDE;s. Nucleotide sequences may be found in a depository of

genetic information, e.g. GenBank, which is a wellknown resource. The inspection of the sequences may be aided by using computer programs that were developed to align several or many sequences, and in so doing highlighting regions of similarity or lack of the same. Many of these are presented and explained in great detail in e.g. Sequence Data Analysis Guidebook /edited by S.R.Swindell, Methods in Molecular Biology vol. 70 (1997), from Humana Press Inc. Totowa, New Jersey.

When sequences have been identified that are unique to a PDE, or respectively shared by several or many PDE;s, oligonucleotide probes based on these sequences may be designed and synthesized. The use of such probes to detect mRNA is well established in the research community, see e.g. Basic DNA and RNA Protocols/edited by A.J.Harwood, Methods in Molecular Biology vol. 58 (1996), from Humana Press Inc. Totowa, New Jersey.

In addition to oligonucleotide probes, mRNA extracted from the tissues and cell types of interest is required, preferably in a form ready to use in Northern analysis. Several companies offer such material, e.g. Invitrogen and Clontech. Briefly, they provide RNA extracted from a great many human and non-human tissues or cell types immobilized on membranes, as an array or size-fractionated.

In a next step, a detectable label needs to be attached to the oligonucleotide probe(s). The label is traditionally in the form of a radioactive isotope, but may to advantage be a chemiluminescent reagent or a fluorescent agent. See e.g. DNA Probes by Keller and Manak (1993), from Macmillan Publishers. Several companies offer reagents to label nucleotide probes, e.g. Ambion (Austin, Texas) and Molecular Probes (Eugene, Oregon).

The actual probing procedure involves contacting the immobilized mRNA (s) with the probe(s), washing away unbound probe(s) and detecting the signal(s) from the probe(s) that bound under the conditions tested, a positive signal indicating that the target(s) of the probe(s) was present in the sample(s) subjected to the test. In its simplest form, the test is "one-to-one", i.e. each sample of mRNA is exposed to each probe. However, it may be advantageous to exploit the sequence hierarchy of the PDE;s, by first probing arrays of mRNA from multiple sources with family-specific probes, then examining first positives with isotype-specific probes, and then examining the secondary positives in detail with very specific probes. One could also multiplex the probing by adding different distinguishable fluorescent labels to the probes, thus obtaining information from several probes in one experiment.

The outcome of the analysis is information regarding the expression pattern(s) of PDE;s.

Based on their expression pattern(s) specific PDE;s are then selected for further study, and genetic probes are constructed.

In general, a genetic probe, i.e. a "GeneX"-GFP fusion or a GFP-"GeneX" fusion, is constructed using PCR with "GeneX"-specific primers followed by a cloning step to fuse "GeneX" in frame with GFP. The fusion may contain a short vector derived sequence between "GeneX" and GFP (e.g. part of a multiple cloning site region in the plasmid) resulting in a peptide linker between "GeneX" and GFP in the resulting fusion protein.

The fusion may be made using polymerase chain reaction techniques, which are common laboratory procedures, see e.g. PCR Protocols/edited by B.A.White, Methods in Molecular Biology vol. 15 (1993), from Humana Press Inc. Totowa, New Jersey.

In more detail, the steps involved include:

- Design of gene-specific primers. Inspection of the sequence of the gene allows design of gene-specific primers to be used in a PCR reaction. Typically, the top-strand primer encompasses the ATG start codon of the gene and the following ca. 20 nucleotides, while the bottom-strand primer encompasses the stop codon and the ca. 20 preceding nucleotides, if the gene is to be fused behind GFP, i.e. a GFP-"GeneX" fusion. If the gene is to be fused in front of GFP, i.e. a "GeneX"-GFP fusion, a stop codon must be avoided. Optionally, the full length sequence of GeneX may not be used in the fusion, but merely the part which localizes and redistributes like GeneX in response to a signal. In addition to gene-specific sequences, the primers contain at least one recognition sequence for a restriction enzyme, to allow subsequent cloning of the PCR product. The sites are chosen so that they are unique in the PCR product and compatible with sites in the cloning vector. Furthermore, it may be necessary to include an exact number of nucleotides between the restriction enzyme site and the gene-specific sequence in order to establish the correct reading frame of the fusion gene and/or a translation initiation consensus sequence. Lastly, the primers always contain a few nucleotides in front of the restriction enzyme site to allow efficient digestion with the enzyme.
- Identifying a source of the gene to be amplified. In order for a PCR reaction to produce a product with gene-specific primers, the gene-sequence must initially be present in the reaction, e.g. in the form of cDNA. The results of the extensive expression analysis performed previously will provide clear information regarding what tissue(s) are useful as source material. cDNA libraries from a great variety of tissues or cell types from various species are commercially available, e.g. from Clontech (Palo Alto), Stratagene (La Jolla) and Invitrogen (San Diego). Many genes are also available in cloned form from The American Type Tissue Collection (Virginia).
- Optimizing the PCR reaction. Several factors are known to influence the efficiency and specificity of a PCR reaction, including the annealing temperature of the primers, the concentration of ions, notably Mg^{2+} and K^+ , present in the reaction, as well as pH of the reaction. If the result of a PCR reaction is deemed unsatisfactory, it might be because

the parameters mentioned above are not optimal. Various annealing temperatures should be tested, e.g. in a PCR machine with a built-in temperature gradient, available from e.g. Stratagene (La Jolla), and/or various buffer compositions should be tried, e.g. the OptiPrime buffer system from Stratagene (La Jolla).

- Cloning the PCR product. The vector into which the amplified gene product will be cloned and fused with GFP will already have been taken into consideration when the primers were designed. When choosing a vector, one should at least consider in which cell types the probe subsequently will be expressed, so that the promoter controlling expression of the probe is compatible with the cells. Most expression vectors also contain one or more selective markers, e.g. conferring resistance to a drug, which is a useful feature when one wants to make stable transfectants. The selective marker should also be compatible with the cells to be used.

The actual cloning of the PCR product should present no difficulty for the person skilled in the art as it typically will be a one-step cloning of a fragment digested with two different restriction enzymes into a vector digested with the same two enzymes. If the cloning proves to be problematic, it may be because the restriction enzymes did not work well with the PCR fragment. In this case one could add longer extensions to the end of the primers to overcome a possible difficulty of digestion close to a fragment end, or one could introduce an intermediate cloning step not based on restriction enzyme digestion. Several companies offer systems for this approach, e.g. Invitrogen (San Diego) and Clontech (Palo Alto).

Once the gene has been cloned and, in the process, fused with the GFP gene, the resulting product, usually a plasmid, should be carefully checked to make sure it is as expected. The most exact test would be to obtain the nucleotide sequence of the fusion-gene.

Once a DNA construct for a probe has been generated, its functionality and usefulness may be tested by subjecting it to the following tests:

- Transfecting it into cells capable of expressing the probe. The fluorescence of the cell is inspected soon after, typically the next day. At this point, two features of cellular fluorescence are noted:

- The intensity should usually be at least as strong as that of unfused GFP in the cells. If it is not, the sequence or quality of the probe-DNA might be faulty, and should be carefully checked.

- The sub-cellular localization is an indication of whether the probe is likely to perform well.

If it localizes as expected for the gene in question, e.g. is excluded from the nucleus, it can immediately go on to a functional test. If the probe is not localized soon after the transfection procedure, it may be because of overexpression at this point in time, as the

cell typically will have taken of very many copies of the plasmid, and localization will occur in time, e.g. within a few weeks, as plasmid copy number and expression level decreases. If localization does not occur after prolonged time, it may be because the fusion to GFP has destroyed a localization function, e.g. masked a protein sequence essential for interaction with its normal cellular anchor-protein. In this case the opposite fusion might work, e.g. if GeneX-GFP does not work, GFP-GeneX might, as two different parts of GeneX will be affected by the proximity to GFP. If this does not work, the proximity of GFP at either end might be a problem, and it could be attempted to increase the distance by incorporating a longer linker between GeneX and GFP in the DNA construct.

If there is no prior knowledge of localization, and no localization is observed, it may be because the probe should not be localized at this point, because such is the nature of the protein fused to GFP. It should then be subjected to a functional test.

In a functional test, the cells expressing the probe are treated with at least one compound known to perturb, usually by activating, the signalling pathway on which the probe is expected to report by redistributing itself within the cell.

If the redistribution is as expected, e.g. if prior knowledge tell that it should translocate from location X to location Y, it has passed the first critical test. In this case it can go on to further characterization and quantification of the response.

If it does not perform as expected, it may be because the cell lacks at least one component of the signalling pathway, e.g. a cell surface receptor, or there is species incompatibility, e.g. if the probe is modelled on sequence information of a human gene product, and the cell is of hamster origin. In both instances one should identify other cell types for the testing process where these potential problems would not apply.

If there is no prior knowledge about the pattern of redistribution, the analysis of the redistribution will have to be done in greater depth to identify what the essential and indicative features are, and when this is clear, it can go on to further characterization and quantification of the response.

If no feature of redistribution can be identified, the problem might be as mentioned above, and the probe should be retested under more optimal cellular conditions.

Libraries for cloning of cDNA libraries in the present discovery plan are naturally related to the target tissues of the projects. For ultimately finding lead compounds useful in the treatment of hyper- and hypotension and erectile dysfunction the cloning libraries should preferably be obtained from one or more of the following tissue or cell types: vascular smooth muscle, vascular smooth muscle from resistance vessels on the arterial side of the vascular system, vascular smooth muscle from capacitance vessels on the venous side of the vascular system, vascular smooth muscle cells from small arteries, arterioles, venules or veins, smooth vascular cells lines such as T/G HA-VSMCA10

and A7r5. For ultimately finding lead compounds useful in the treatment of jet-lag the cloning libraries should preferably be obtained neurons from the brain.

The cells should always be of animal origin, most likely of mammalian origin and preferably of human origin. The cells could be derived from normal tissue or from tissue of an individual animal having a disease or condition of interest for the project. The cells may also be a mortal or immortalised cell line where the initial cell clone has been derived from a tissue or cell type as described above. Depending on the discovery project the cells of interest for screening assays will vary but may be chosen from the above mentioned categories.

Once a genetic construct containing the protein of interest and the luminophore, from here on referred to as "the original fluorescent probe", has been transfected into a relevant cell type, as described above under 'preferred cell types for cloning libraries' the cells are monitored for the appearance of spatially distributed or randomly distributed intracellular fluorescence. Based on prior knowledge regarding the distribution of the actual protein different patterns can be expected. If for example previous studies have found the protein associated only with the particulate fraction of fractionated cells, it can be expected to find a spatial distribution of the original fluorescent probe to the plasma membrane, internal membrane/organelle structures or structural cytoplasmic elements such as microtubules and microfilaments. If on the other hand previous studies report that the protein has been found mostly in the soluble fraction of fractionated cells one can expect to find a homogenous or nonhomogenous distribution of the original fluorescent probe throughout the cytoplasm and perhaps also in the nucleus. For proteins where previous studies have found a mixed localisation to both the particulate and soluble fraction of fractionated cells any mixture in the two distribution patterns mentioned above for the original fluorescent probe can be expected. For proteins where no prior knowledge is at hand a simple cell fractionation and Western Blotting can be made, one can use immunohistochemistry of fixed cells of relevance or one can decide to rely on the distribution observed for the original fluorescent probe. At this stage of the project, a normal distribution pattern of the original fluorescent probe may be established after such studies as outlined above. The effects of physiologically important and relevant cellular activation on the distributed pattern of the original fluorescent probe is also established. It will also become evident if the pattern of distribution changes, i.e. if a redistribution of the original fluorescent probe occurs as a consequence of applying a physiologically important and relevant influence.

By the methods described in appendix A it may be possible to find chemical entities which can interfere with the protein-protein interactions that occur amongst PDE:s and

their anchoring/regulating partners, and thereby interfere with the effectiveness of a PDE isoform's action within its cellular environment. This strategy will have different effects (and require slightly different discovery methods) depending on the nature of the interaction. The possibilities are as follows:

- 1) A PDE isoform is permanently located at its targeting point, and either remains permanently active there, or its activity is modulated in some way by post-translational modification such as phosphorylation or by binding of modulators to non-catalytic regulatory sites. Dislocation from the targeting site will remove the PDE from a localised site of action, and may also lead to inactivation of its inherent catalytic activity.
- 2) A PDE isoform is permanently located at its targeting point, and remains inactive there until its activity is modulated in some way by post-translational modification, such as phosphorylation or by binding of modulators to non-catalytic regulatory sites. Dislocation from the targeting site will remove the PDE from a localised site of action, and may also lead to activation of its inherent catalytic activity, albeit away from its original anchoring site.
- 3) A PDE isoform is inactive in its unattached or untargeted form, and when activated (as described in "1" above), or partially activated, it redistributes within the cell and becomes attached to its targeting site, its activity being restricted to the anchoring site and possibly enhanced by interaction with the anchoring protein or some associated factor, or at some later time inhibited by the anchoring protein or an associated regulatory factor. Any agent which prevents association of the PDE with its anchoring or targeting site will prevent it from locating to the preferred site of action, and may also prevent the PDE from becoming fully activated by the appropriate stimulus whilst in the untargeted state.
- 4) A PDE isoform is active in its unattached or untargeted form, and when inactivated (as described in "1" above), or partially inactivated, it redistributes within the cell and becomes attached to its targeting site, whereby its activity is inhibited by interaction with the anchoring protein or an associated regulatory factor. Subsequent stimuli may then activate and release the PDE. Any agent which prevents association of the PDE with its anchoring or targeting site will prevent it from relocating to the anchoring position, and may also prevent the PDE from ever being inactivated. In addition, if the PDE cannot target to its anchoring site, it may not be possible subsequently to activate the PDE in the appropriate way in the untargeted state.

When a specific subcellular distribution of a GFP-based PDE probe has been identified, it may be advantageous to narrow down which part of the PDE is responsible for this effect. The advantage is twofold: It may suggest the design of peptide leads, and it may eventually aid in defining the binding partner. Knowledge of both partners involved in

specific binding may aid in the selection of compound libraries to screen for inhibition of the specific binding.

To identify the region of the PDE involved in specific binding, one may make GFP-based fusions with progressively shorter parts of the PDE, and examine the cellular distribution of these constructs. If there is prior knowledge of functional domains, one may start with the domain believed to confer specific binding to a subcellular structure. The generation of constructs to test may consist of selecting a particular part of the PDE to fuse to GFP, or it may involve the generation of in-frame deletions in the PDE part of the fusion. Both approaches have been widely used in molecular genetic studies.

When a region has been identified that appears responsible for conferring a specific subcellular distribution upon a PDE, the amino acid residues most important for this trait may be identified by a more detailed analysis, e.g. substituting them one by one with e.g. an alanine residue, a so called Ala-scan, which also has been used extensively in molecular genetic studies.

To identify the identity of the cellular protein partaking in the specific distribution of the PDE, one may exploit the knowledge about the region of the PDE responsible for the subcellular distribution. e.g. one may use the region of the PDE as bait in a genetic two hybrid screen to pull out its binding partner. Several companies offer two hybrid systems, e.g. Life Technologies.

The knowledge about the normal distribution of the original fluorescent probe is used to establish which part or which parts of the terminal (or entire) amino-acid sequence that is important for the attachment of this fluorescent probe to subcellular structures, giving it its specific spatially distributed pattern in the cell or cells, when such a pattern has been established as the normal distribution of this fluorescent probe. This may be accomplished by creating new fluorescent probes where a systematic deletion of short N- or C-terminal or internal sequences (number of DNA bases) of the original fluorescent probe are made. These new shorter variants of the of the original fluorescent probe construct are transfected into the cells of interest and then the cells are examined for spatial distribution of the new fluorescent probes as described above for the original fluorescent probe. In those cells where the new fluorescent probe distribution pattern is different from the original fluorescent probe distribution pattern it is evident that part of the, or the entire, targeting sequence has been deleted. The DNA- or amino-acid sequence of the missing part therefore contains the structural information necessary for association of the original fluorescent probe with its intracellular binding sites.

Peptides for inhibition of the established normal distribution of the original fluorescent probe are made according to the following hypothesis, that the deduced targeting

sequence, or sequences, in the original fluorescent probe amino-acid sequence are the important sequences for the actual spatial distribution of the original fluorescent probe in intact living cells, is tested. This is done by producing peptides of identical amino-acid sequence as the deduced targeting sequence or parts thereof and introducing them into the cytoplasm, either by microinjection or transient or permanent permeabilisation, of cells containing the original fluorescent probe and thereafter monitoring the spatial distribution of the original fluorescent probe in the cells. If the deduced targeting sequence or sequences are of importance for the actual spatial distribution of the original fluorescent probe in intact living cells, the introduced peptides will self-associate with the anchoring sites for the original fluorescent probe and thereby disrupt the normal distribution of the original fluorescent probe. In order to have this effect, the introduction of the peptides should change the original distribution pattern so that a decrease in fluorescence of 10% or more, compared to the pattern before their introduction, can be detected. This is done by observing the same cells before and after administration of the peptides. When peptides that fulfil this criterion have been found they are called 'peptide leads' and will hereafter be referred to using this expression. These peptide leads can now be used as a basis for the design of organic molecules that can be used eventually to disrupt the spatial distribution of the original fluorescent probe but also as control compounds in screening assays.

In parallel to the above mentioned procedure wherein peptide leads are defined, the distribution pattern found for the original fluorescent probe is compared to the naturally occurring spatial distribution of the protein on which the original fluorescent probe is based. This may be accomplished by fixation of primary cells separated or within the tissue of interest and fixation of cells that contain the original fluorescent probe. Thereafter the protein is stained using ordinary immunocytochemical or immunohistochemical methods and the spatial distribution revealed by this staining procedure is compared to the spatial distribution of the original fluorescent probe. It is desirable, but not required, that a high degree of correlation between the two patterns obtained in this step can be observed.

The establishment of a primary screening assay is normally done by making use of the cells of interest containing the original fluorescent probe as the basis for a screening assay. Depending on the knowledge acquired about the behaviour of the original fluorescent probe when subjecting the cells to physiologically relevant influences the assay procedure can be chosen: 1. If the fluorescent probe normally is targeted to specific sites and stay associated with these sites during stimulation of the intracellular pathway the assay should preferably be designed to detect dislocation of the original fluorescent probe from the targeting sites in mechanically intact or permeabilised living

cells. This is an assay where the dislocation can be detected within minutes after application of an influence and the time frame for the detection and time for exposing the cells to an influence should be chosen to match this. 2. If the desire is to disrupt the actual targeting event rather than dislocate already targeted fluorescent probe the influence may need hours to produce a detectable response. The actual measurement, still of a change in the fluorescence or luminescence distribution pattern compared to the normal distribution pattern for the original fluorescent probe, may be made at two time points; before and after the influence has exerted any effect it may have. This is an assay where the effect of an influence may require several hours to produce a detectable response and the time frame for the detection and time for exposing the cells to an influence should be chosen to match this. 3. If the fluorescent probe normally redistributes between two intracellular sites upon activation of the intracellular pathway one may either want to disrupt the initial targeting or dislocate the original fluorescent probe from its initial or resting anchoring site. In this case procedure no. 1 above may be used. If the desire instead is to inhibit the association of the original fluorescent probe with the site it redistributes to during activation of the intracellular pathway the targeting sequence of this site should be in focus for the lead peptide generation. This is an assay where the redistribution may be detected within minutes after application of an influence and the time frame for the detection and time for exposing the cells to an influence should be chosen to match this. Furthermore, any influence applied to inhibit the targeting of the original fluorescent probe upon its redistribution may need to be added to the cells before activation of the intracellular pathway.

While the original fluorescent probe and peptide leads will be used in the actual primary screening assay, it is also desirable to have a counterscreen or counterscreens directed at protein isoforms that one does not wish to affect. In order to accomplish this, constructs are made for new fluorescent probes encoding the protein isoforms tagged with GFP. These constructs are subsequently transfected into the cells of interest. When the new fluorescent probes are expressed in the cells, some of the cells are chosen as the basis for new cell lines that can be used in the counterscreen or counterscreens.

Suitable probes for this purpose comprise DNA constructs encoding fusion polypeptides comprising forms of PDE 1, PDE 2, PDE 5, PDE 6 or PDE 9 and GFP. In a preferred embodiment the DNA constructs will encode fusion polypeptides comprising isoforms of mPDE 5 and GFP.

It may be desirable to establish peptide leads that more specifically dislocate the desired isoform of the protein of interest compared to other isoforms of the same protein. In such a case, both the cell lines established for the primary screen and the counterscreen

or counterscreens are used. The peptide leads are introduced into the cells as described above and the changes in spatial distribution of the original and counterscreen fluorescent probes are quantified and dose-response relationships are established for each lead peptide. Thereafter the dose-response relationships are compared. A peptide lead is considered specific for the original fluorescent probe if the dose of the peptide required to dislocate at least 10% of the fluorescent probes in the counterscreen or counterscreens are at least two times higher than the dose required to dislocate 10% of the original fluorescent probe. The lead peptides with the biggest dose difference when comparing the primary and the counterscreen dose-response relationships are chosen as the basis for the next step in the discovery project.

In one embodiment the primary screening assay and counterscreen or counterscreens are used to define specificity of the peptide leads by using a procedure that compares their ability to cause a dislocation, disruption of targeting or inhibition of redistribution of the original fluorescent probe in the primary screening assay to their ability to cause a dislocation, disruption of targeting or inhibition of redistribution of the new fluorescent probes in the counterscreen or counterscreens.

In a preferred embodiment the dose of a peptide lead required to cause a quantified dislocation, disruption of targeting or inhibition of redistribution of the original fluorescent probe of at least 10% in the primary screening assay is 50% or less of the dose required to cause a quantified dislocation, disruption of targeting or inhibition of redistribution of the new fluorescent probes of at least 10% in the counterscreen or counterscreens.

The invention provides for a specificity index which may be constructed describing a numerical relationship, with the primary screening assay result first, of the dose required to produce half-maximal effect in the primary assay compared to the dose required to produce half-maximal effect in the counterscreen or counterscreens.

In one embodiment the peptide leads chosen for further use in the discovery project have a specificity index of 1 to 2.

In another embodiment the peptide leads chosen for further use in the discovery project have a specificity index between 1 to 2 and 1 to 10.

In a further embodiment the peptide leads chosen for further use in the discovery project have a specificity index between 1 to 11 and 1 to 100.

In yet a further preferred embodiment the peptide leads chosen for further use in the discovery project have a specificity index better than 1 to 100.

The lead peptides may be used to create and select libraries of small organic molecules that can be useful in screening assays to find bioactive substances useful as drugs to treat the condition or disease of interest for the project. In order to accomplish this, the

amino-acid sequence information and other structural information about the lead peptide or peptides is used to extract information useful for finding and/or defining and synthesising bioactive organic molecules that can mimic the effect of the lead peptides on the normal spatial distribution pattern of the original fluorescent probe. Peptide leads selected by the discovery project are used to design and assemble compound libraries based on the structural and chemical information inherent in the lead peptides using prior chemical knowledge and computational chemistry approaches so that the compounds have a structure that give them the ability to interact with or bind to the targeting sequence of mPDE 5, thereafter testing the compound libraries at a concentration of 10 or 100 micromolar of each compound in the primary screening assay.

When the libraries of compounds have been defined and are at hand it is time to initiate primary screening. In this procedure, cells containing the original fluorescent probe are contacted with the compounds. The compounds are all tested at just one or a few concentrations, typically 10 and 100 micromolar, in a highly parallel fashion using a quantitative fluorescence redistribution assay. Compounds that cause a change in the quantitated response (the response scale defined by the range 0 (no change in redistribution) – 100%) of the assay by more than a predetermined value, typically between 10 and 100%, are considered to be "primary hits". The primary hits are then further characterised: 1. for potency by establishing a dose-response relationship compared to the lead peptide(s) using the primary screening assay 2. for selectivity by establishing a dose-response relationship in the counterscreen or counterscreens. Primary hits that have low potency, typically when the half-maximal effect of the compound in the primary assay is achieved at a concentration of the compound between 10 and 100 micromolar, may not need testing in the counterscreen or counterscreens since the likelihood that they will be used beyond this step in the discovery project is small. Primary hits that have equal or lower potency in the primary screening assay compared to the counterscreen or counterscreens are regarded as non-selective and the likelihood that they will be used beyond this step in the discovery project is small. Primary hits that display some degree of selectivity, typically half maximal effect in the primary screening assay at a concentration 50% or less of the concentration that gives half maximal effect in the counterscreen or counterscreens are considered interesting as the basis for further chemical synthesis or construction of new libraries of compounds and will hereafter be referred to as "primary lead compounds".

Compounds that cause a change in the quantitated response, with a response scale from 0 to 100% based on the absence of a response and the maximal response observed with the peptide leads in the primary screening assay, of the assay by more than a predetermined value are selected and called "primary hits".

In one embodiment the predetermined value is 10%.

In another embodiment the predetermined value is 50%.

In yet another embodiment the predetermined value is 70%.

In one embodiment the primary hits are further characterised for potency (as defined herein) and maximal effect by establishing a dose-response relationship (as defined herein) and comparing that to the effects of the lead peptides using the primary screening assay and for selectivity (as defined herein) by establishing a dose-response relationship in the counterscreen or counterscreens.

Primary hits may be deselected by the discovery project when they display a half-maximal potency at a dose corresponding to a concentration of more than 10 micromolar or because they display a selectivity index less than 1 to 2.

Primary hits may be selected by the discovery project when they display a half-maximal potency at a dose corresponding to a concentration of 10 micromolar or less or because they display a selectivity index higher than 1 to 2, the compounds hereafter also referred to as "primary lead compounds".

To further improve the possibilities of finding bioactive compounds with desirable properties for treatment of the diseases or conditions of interest to the project one may at this stage build a Structure-Activity Relationship (SAR) by iterations of compound library composition and screening to define drug candidate leads. The primary lead compounds are here used to provide chemical structural information that can be used as the basis for composition or chemical synthesis of new, directed, compound libraries. By systematic chemical modification of part of the structure of one or more primary lead compounds new libraries are assembled. These new libraries of compounds are also investigated using the primary screening assay and counterscreen or counterscreens. Preferably, dose-response relationships are recorded for each chemical modification of the primary lead compound and compared to the primary lead compound itself. Thereby SAR is established. Among the new compounds, the ones that in this step has the best combination of potency and specificity are chosen either as the basis for a new round of compound library synthesis or composition or, as the final step of the SAR building process, as compounds that will be further for actual pharmacological effects in assay systems and animals that are relevant to the underlying physiological and pathophysiological processes of interest to the project. The latter compounds will hereafter be referred to as "drug candidate leads".

In one embodiment drug candidate leads have a half-maximal potency at a dose corresponding to a concentration of less than 1 micromolar and a selectivity index higher than 1 to 2.

In one embodiment the drug candidate leads have a half-maximal potency at a dose corresponding to a concentration of less than 1 micromolar and a selectivity index higher than 1 to 10.

In one embodiment the drug candidate leads have a half-maximal potency at a dose corresponding to a concentration of less than 1 micromolar and a selectivity index higher than 1 to 100.

In one embodiment the drug candidate leads have a half-maximal potency at a dose corresponding to a concentration of less than 0,1 micromolar and a selectivity index higher than 1 to 2.

In a preferred embodiment the drug candidate leads have a half-maximal potency at a dose corresponding to a concentration of less than 0,1 micromolar and a selectivity index higher than 1 to 10.

In another preferred embodiment the drug candidate leads have a half-maximal potency at a dose corresponding to a concentration of less than 0,1 micromolar and a selectivity index higher than 1 to 100.

To validate the drug candidate leads *in vitro* is often preferred at this stage. In this case, drug candidate leads may be further characterised in tissue based, cell based and biochemical assays for efficacy and toxicity. There are many ways to test efficacy of a drug candidate lead. Preferably, the drug candidate lead is tested in assay systems with high relevance to the underlying physiological and patophysiological processes involved in the pathogenesis and patophysiology of the disease or condition of interest to the project. Likewise, the drug candidate leads are tested for toxic effects, preferably testing for genetic effects (influence on the integrity and arrangement of DNA), metabolic effects (influence on cellular metabolic processes) and cytotoxic effects (influence on cell integrity and organelle integrity). There is a high likelihood that drug candidate leads, that do not show appropriate efficacy or that display toxicity will not be used beyond this step in the discovery project because it is expected that such compounds are less suitable as actual drugs to be used in an animal.

In one embodiment drug candidate leads chosen by the discovery project are tested *in vitro* for efficacy (as defined herein), in assay systems with high degree of relevance to the underlying physiological and patophysiological processes involved in hypotension, and for toxicity, preferably testing for genetic, metabolic and cytotoxic effects, whereafter the drug candidate leads that display the best efficacy and the least, or no, indications of toxicity are chosen to be the candidates that will enter testing in animals.

In another embodiment drug candidate leads chosen by the discovery project are tested *in vitro* for efficacy, in assay systems with high degree of relevance to the underlying physiological and patophysiological processes involved in hypertension, and for

toxicity, preferably testing for genetic, metabolic and cytotoxic effects, whereafter the drug candidate leads that display the best efficacy and the least, or no, indications of toxicity are chosen to be the candidates that will enter testing in animals.

In another embodiment drug candidate leads chosen by the discovery project are tested *in vitro* for efficacy, in assay systems with high degree of relevance to the underlying physiological and pathophysiological processes involved in jet-lag, and for toxicity, preferably testing for genetic, metabolic and cytotoxic effects, whereafter the drug candidate leads that display the best efficacy and the least, or no, indications of toxicity are chosen to be the candidates that will enter testing in animals.

In another embodiment drug candidate leads chosen by the discovery project are tested *in vitro* for efficacy, in assay systems with high degree of relevance to the underlying physiological and pathophysiological processes involved in erectile dysfunction, and for toxicity, preferably testing for genetic, metabolic and cytotoxic effects, whereafter the drug candidate leads that display the best efficacy and the least, or no, indications of toxicity are chosen to be the candidates that will enter testing in animals.

Finally, it is mostly likely preferred to do an *in vivo* validation of the drug candidate leads. In this case the drug candidate leads are tested for toxic and unwanted effects in animals such as mice and rats. The drug candidate leads are also tested for efficacy in animals that have a disease or condition with high degree of relevance to the disease or condition of interest to the project. The drug candidate leads may also be tested for efficacy in animals which have been treated in a way that make them experience a disease or condition with high degree of relevance to the disease or condition of interest to the project. Drug candidate leads that display efficacy in one or more of such animal tests and that does not display any apparent toxicity at a dosage level, preferably 2 –10 times higher than the level that gives satisfactory efficacy are chosen to be the final drug candidates that should be considered for further animal testing and initial testing in humans. These compounds are hereafter referred to as “discovery project leads”.

In one embodiment drug candidate leads chosen by the discovery project are tested for efficacy, in healthy animals and animals with a condition with high degree of relevance to the underlying physiological and pathophysiological processes involved in hypotension, and for toxicity and unwanted side effects, after which the drug candidate leads that display the best efficacy and the least, or no, indications of toxicity or unwanted side effects are chosen to be the candidates, called discovery project leads, that will enter further testing in animals and testing in humans.

In one embodiment drug candidate leads chosen by the discovery project are tested for efficacy, in healthy animals and animals with a condition with high degree of relevance to the underlying physiological and pathophysiological processes involved in

hypertension, and for toxicity and unwanted side effects, after which the drug candidate leads that display the best efficacy and the least, or no, indications of toxicity or unwanted side effects are chosen to be the candidates, called discovery project leads, that will enter further testing in animals and testing in humans.

In one embodiment drug candidate leads chosen by the discovery project are tested for efficacy, in healthy animals and animals with a condition with high degree of relevance to the underlying physiological and pathophysiological processes involved in jet-lag, and for toxicity and unwanted side effects, after which the drug candidate leads that display the best efficacy and the least, or no, indications of toxicity or unwanted side effects are chosen to be the candidates, called discovery project leads, that will enter further testing in animals and testing in humans.

In one embodiment drug candidate leads chosen by the discovery project are tested for efficacy, in healthy animals and animals with a condition with high degree of relevance to the underlying physiological and pathophysiological processes involved in erectile dysfunction, and for toxicity and unwanted side effects, after which the drug candidate leads that display the best efficacy and the least, or no, indications of toxicity or unwanted side effects are chosen to be the candidates, called discovery project leads, that will enter further testing in animals and testing in humans.

The administration route of any of the compounds of the invention may be of any suitable route which leads to a concentration in the blood corresponding to a therapeutic concentration by the oral route, the parenteral route, the cutaneous route, the nasal route, the rectal route, the vaginal route and the ocular route. It should be clear to a person skilled in the art that the administration route is dependant on the compound in question, particularly, the choice of administration route depends on the physico-chemical properties of the compound together with the age and weight of the patient and on the particular disease and the severity of the same.

The compounds of the invention may be contained in any appropriate amount in a pharmaceutical composition, and are generally contained in an amount of about 1-95% by weight of the total weight of the composition. The composition may be in form of, e.g., tablets, capsules, pills, powders, granulates, suspensions, emulsions, solutions, gels including hydrogels, pastes, ointments, creams, plasters, drenches, delivery devices, suppositories, enemas, injectables, implants, sprays, aerosols and in other suitable form. The pharmaceutical compositions may be formulated according to conventional pharmaceutical practice, see, e.g., "Remington's Pharmaceutical Sciences" and "Encyclopedia of Pharmaceutical Technology".

Pharmaceutical compositions according to the present invention may be formulated to release the active compound substantially immediately upon administration or at any substantially predetermined time or time period after administration. The latter type of

compositions are generally known as controlled release formulations. Controlled release formulations may also be denoted "sustained release", "prolonged release", "programmed release", "time release", "rate-controlled" and/or "targeted release" formulations.

In the present context every pharmaceutical composition is an actual drug delivery system, since upon administration it presents the active drug substance to the body of the organism.

The compounds of the invention are preferably administered in an amount of about 0.1-30 mg per kg body weight per day, such as about 0.5-15 mg per kg body weight per day. The compound in question may be administered orally in the form of tablets, capsules, elixirs or syrups, or rectally in the form of suppositories. Parenteral administration of the compounds of the invention, is suitably performed in the form of saline solutions of the compounds or with the compound incorporated into liposomes. In cases where the compound in itself is not sufficiently soluble to be dissolved, an acid addition salt of a basic compound can be used, or a solubilizer such as ethanol can be applied.

Oral administration. For compositions adapted for oral administration for systemic use, the dosage is normally 1 mg to 1 g per dose administered 1-4 times daily for 1 week, 12 months or even lifelong depending on the disease to be treated.

Rectal administration. For compositions adapted for rectal a somewhat higher amount of compound is usually preferred, i.e. from approximately 1 mg to 100 mg per kg body weight per day.

Parenteral administration. For parenteral administration a dose of about 0.1 mg to about 50 mg per kg body weight per day is convenient. For intravenous administration a dose of about 0.1 mg to about 20 mg per kg body weight per day. For intraarticular administration a dose of about 0.1 mg to about 20 mg per kg body weight per day is usually preferable. For parenteral administration in general, a solution in an aqueous medium of 0.5-2% or more of the active ingredients may be employed.

Percutaneous administration. For topical administration on the skin a dose of about 1 mg to about 5 g administered 1-10 times daily is usually preferable.

EXAMPLES

Probes for detection of PDE5 dislocation:

These are specific PDE5 variants fused to a GFP. Currently only one main human variant is known (GenBank Acc.nos. AJ004865 and D89094).

Inspection of the scientific literature indicates that the catalytic domain is contained in the C-terminal part of the protein, so potential targeting sequences of PDE5 may be located in the N-terminal part. To best preserve the normal distribution of PDE5, the first fusion is made between the C-terminus of the PDE5 species and the N-terminal of the GFP.

To construct the PDE5-GFP fusions, PDE5 sequences are amplified using PCR according to standard protocols with the specific primers listed below. The PCR product is digested with restriction enzymes EcoR1 and Acc65I, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with EcoR1 and Acc65I. This produces a PDE5-EGFP fusion under the control of a CMV promoter (SEQ.ID.NO 1&2).

The top primer includes specific sequences following the ATG, a Kozak sequence, and a cloning site (EcoR1). The bottom primer includes specific C-terminal sequences minus the stop codon, an Acc65I cloning site, and two extra nucleotides to preserve the reading frame in EGFP-N1.

PDE5-top (SEQ.ID.NO. 3):
5'-GTGAATTCAACCATGGAGCGGGCC-3'

PDE5-bottom (SEQ.ID.NO. 4):
5'-GTGGTACCCAGTCCGCTTGGCC

The resulting plasmids are transfected into a suitable cell line, e.g. MVLEC. The subcellular distribution of the probes is examined carefully by fluorescence microscopy, both under resting conditions, and upon elevation of cGMP, e.g. by activation of cyclase with NO or nitroprusside, which may or may not have an effect on the normal distribution.

REFERENCES

- Ashida, S., Sakuma, K.. (1992) Adv. Second Messenger Phosphoprotein Res. 25:229-239.
- Beavo, J.A., (1995) Physiol. Rev 75:725-748
- Benner, B.M., Ballermann, B.J., Gunning, M.E., Zeidel, M.L. (1990) Physiol. Rev. 70:665-699
- Biel, M., Zong, X., Distler, M., Bosse, E., Klugbauer, N., Murakami, M., Flockerzi, V. and Hofmann, F. (1994) Comment in: Proc Natl Acad Sci USA. 91: 3505-3509.
- Bolger *et al.*, (1997) Biochem. J. 328:539-548.
- Bolger, G.B., McPhee, I., Houslay, M.D. (1996) J. Biol. Chem. 271:1065-1071
- Burns, F., Rodger, I.W. and Pyne, N.J. (1992) Biochem. J. 283 (Pt 2): 487-491.
- Chiu, P.J., Vemulapalli, S., Chintala, M., Kurowski, S., Tetzloff, G.G., Brown, A.D. and Sybertz, E.J. (1997) Naunyn-schmiedebergs Archives of Pharmacology, 355: 463-469.
- Drewett, J.G. and Garbers, D.L. (1994) Endocrine Reviews, 15: 135-162.
- Faux, M.C., Scott, J.D., (1996) Cell 85:9-12.
- Fisher, D.A., Smith, J.F., Pillar, J.S., Stdenis, S.H. and Cheng, J.B. (1998) J. Biol. Chem. 273: 15559-15564.
- Hofmann, F., Dostmann, W., Keilbach, A., Landgraf, W. and Ruth, P. (1992) Biochimica Et Biophysica Acta, 1135: 51-60.
- Holliday, L.S., Dean, A.D., Lin, R.H., Greenwald, J.E. and Gluck, S.L. (1997) Am. J. Physiol. 272: F283-F291
- Houslay, M.D., Milligan, G. (1997) Trends in Biochem. Sci. 22:217-224.
- Houslay, M.D., Sullivan, M., Bolger, G.B. (1998) Adv. Pharmacol. 44:225-242
- Hughes B. *et al.*, (1997) Drug Discov. Today 2:89-101.
- Keilbach, A., Ruth, P. and Hofmann, F. (1992) Eur. J. Biochem. 208: 467-473.
- Light, D.B., Corbin, J.D. and Stanton, B.A. (1990) Nature, 344: 336-339.
- Lincoln, T.M. and Cornwell, T.L. (1993) Faseb Journal, 7: 328-338.
- Lincoln, T.M., Komalavilas, P. and Cornwell, T.L. (1994) Hypertension, 23: 1141-1147.
- Liu, C., Ding, J.M., Faiman, L.E. and Gillette, M.U. (1997) J. Neuroscience, 17: 659-666.
- Lochhead, A., Nekrasova, E., Arshavsky, V.Y. and Pyne, N.J. (1997) J. Biol. Chem. 272: 18397-18403.
- MacFarland, R.T., Zelus, B.D., Beavo, J.A., (1991) J. Biol. Chem. 266:136-142.
- MacMillan-Crow, L.A. and Lincoln, T.M. (1994) Biochemistry, 33: 8035-8043.

- Mathur, A., Golombek, D.A. and Ralph, M.R. (1996) *Am. J. Physiol.* 270: R1031-R1036
- McPhee, I., *et al.*, (1995) *Biochem. J.* 310:965-974.
- O'Connell, J.C., *et al.*, (1996) *Biochem. J.* 318:255-262.
- Pawson, T., Scott, J.D. (1997) *Science* 278:2075-2080.
- Perry, M.J., Higgs, G.A. (1998) *Curr. Opin. Chem. Biol.* 2:472-481.
- Pooley, L., Shakur, Y., Rena, G., Houslay, M.D. (1997) *Bioch. J.* 321:177-185.
- Pyne, N.J., Arshavsky, V. and Lochhead, A. (1996) *Biochem. Soc. Trans.* 24: 1019-1022.
- Pyne, N.J., Arshavsky, V., Lochhead, A., (1996) *Biochem. Soc. Trans.* 24:1019-1022.
- Pyne, N.J., Cooper, M.E. and Houslay, M.D. (1987) *Biochemical Journal*, 242: 33-42.
- Raeburn, D. and Karlsson, J.A. (1993) *J. Pharm. Exp. Ther.* 267: 1147-1152.
- Scotland *et al.*, (1998) *Methods - A companion to Meth. Enzymology* 14:65-79.
- Sette, C., Conti, M., (1996) *J. Biol. Chem.* 271:16526-16534.
- Shakur, Y. *et al.*, (1995) *Biochem. J.* 310:965-974.
- Soderling, S.H., Bayuga, S.J. and Beavo, J.A. (1998) *J. Biol. Chem.* 273: 15553-15558.
- Teixeira, M.M., Gristwood, R.W., Cooper, N., Hellewell, P.G. (1997) *Trends in Pharm. Science* 18:164-171.
- Thomas, M.K., Francis, S.H. and Corbin, J.D. (1990) *J. Biol. Chem.* 265: 14971-14978.
- Vaandrager, A.B. and de Jonge, H.R. (1996) *Mol. Cell. Biochem.* 157: 23-30.
- Vemulapalli, S., Watkins, R.W., Chintala, M., Davis, H., Ahn, H.S., Fawzi, A., Tulshian, D., Chiu, P., Chatterjee, M., Lin, C.C. and Sybertz, E.J. (1996) *J. Cardiovasc. Pharm.* 28: 862-869.

CLAIMS

1. A method for preventing or treating, in an animal in need thereof, an adverse condition which may be reduced or abolished by modulating the activity of one or more cyclic nucleotide phosphodiesterases having the ability to cleave cyclic GMP, the method comprising modulating the specific effectiveness of the cyclic nucleotide phosphodiesterase by modulating their spatial distribution within cells of the animal.
2. A method according to claim 1, wherein the cyclic nucleotide phosphodiesterase is selected from the group consisting of PDE 1, PDE 2, PDE 5, PDE 6, and PDE 9.
3. A method according to claim 1 or 2, wherein the cyclic nucleotide phosphodiesterase is a PDE 5 species.
4. A method according to claim 3, wherein the PDE 5 species is a mPDE 5 species.
5. A method according to any of claims 1-4, wherein the animal is a mammal.
6. A method according to claim 5, wherein the mammal is a human being.
7. A method according to any of claims 1-6, wherein the modulation of the specific effectiveness of the cyclic nucleotide phosphodiesterase is a dislocation from a native location within the cell.
8. A method according to any of claims 1-6, wherein the modulation of the specific effectiveness of the cyclic nucleotide phosphodiesterase involves a disruption of the targeting of the cyclic nucleotide phosphodiesterase to a native location within the cell.
9. A method according to any of claims 1-6, wherein the modulation of the specific effectiveness of the cyclic nucleotide phosphodiesterase involves interference with the redistribution of the cyclic nucleotide phosphodiesterase, the redistribution being associated with an increase or a decrease in the specific effectiveness of the cyclic nucleotide phosphodiesterase.
10. A method according to any of claims 1-9, wherein the adverse condition is hypotension.

14. A method according to any of claims 1-9, wherein the adverse condition is hypertension.
15. A method according to any of claims 1-9, wherein the adverse condition is cardiac failure or thrombosis.
16. A method according to any of claims 1-9, wherein the adverse condition is erectile dysfunction.
18. A method according to any of claims 1-9, wherein the adverse condition is jet lag.
19. A method according to any of claims 1-9, wherein the adverse condition is oedema caused by inflammation.
20. A method according to any of the preceding claims, wherein the modulation of the specific effectiveness of the cyclic nucleotide phosphodiesterase is performed by exposing cells, in the animal in which dislocation, disruption of targeting, or interference with redistribution of a cyclic nucleotide phosphodiesterase may take place, to the influence of a substance which modulates the spatial distribution of the cyclic nucleotide phosphodiesterase in the cells.
21. A method according to claim 20, wherein the substance is one which, in a quantitative fluorescence redistribution assay designed to monitor dislocation of cyclic nucleotide phosphodiesterase, causes dislocation of at least 10% of otherwise natively located cyclic nucleotide phosphodiesterase within the cell at a concentration of the substance of 100 micromolar.
22. A method according to claim 21, wherein at least 50% of otherwise natively located cyclic nucleotide phosphodiesterase is dislocated within the cell at a concentration of the substance of 100 micromolar..
23. A method according to claim 21, wherein at least 70% of otherwise natively located cyclic nucleotide phosphodiesterase is dislocated within the cell at a concentration of the substance of 100 micromolar.
24. A method according to claim 21, wherein at least 90% of otherwise natively located cyclic nucleotide phosphodiesterase is dislocated within the cell at a concentration of the substance of 100 micromolar.

25. A method according to claim 20, wherein the substance is one which, in a quantitative fluorescence redistribution assay, designed to monitor targeting of cyclic nucleotide phosphodiesterase, reduces targeting of the cyclic nucleotide phosphodiesterase to its native location within the cell by at least 10% at a concentration of the substance of 100 micromolar.

26. A method according to claim 25, wherein the substance reduces targeting of the cyclic nucleotide phosphodiesterase to its native location within the cell by at least 50% at a concentration of the substance of 100 micromolar.

27. A method according to claim 25, wherein the substance reduces targeting of the cyclic nucleotide phosphodiesterase to its native location within the cell by at least 70% at a concentration of the substance of 100 micromolar.

28. A method according to claim 25, wherein the substance reduces targeting of the cyclic nucleotide phosphodiesterase to its native location within the cell by at least 90% at a concentration of the substance of 100 micromolar.

29. A method according to claim 20, wherein the substance is one which, in a quantitative fluorescence redistribution assay, designed to monitor changes in redistribution caused by an influence, causes a reduction in the induced redistribution by at least 10% of the normal maximum redistribution at a concentration of the substance of 100 micromolar.

30. A method according to claim 29, wherein the substance causes a reduction in the induced redistribution of the cyclic nucleotide phosphodiesterase by at least 50% of the normal maximum redistribution at a concentration of the substance of 100 micromolar.

31. A method according to claim 29, wherein the substance causes a reduction in the induced redistribution of the cyclic nucleotide phosphodiesterase by at least 70% of the normal maximum redistribution at a concentration of the substance of 100 micromolar.

32. A method according to claim 29, wherein the substance causes a reduction in the induced redistribution of the cyclic nucleotide phosphodiesterase by at least 90% of the normal maximum redistribution at a concentration of the substance of 100 micromolar.

33. A method according to any of claims 20-32, wherein the substance is an organic compound having a molecular weight of at the most 1200 Da.

34. A method according to any of claims 20-32, wherein the substance is an organic compound having a molecular weight of at the most 900 Da.

35. A method according to any of claims 20-32, wherein the substance is an organic compound having a molecular weight of at the most 600 Da.

36. A method according to any of claims 20-32, wherein the substance is an organic compound having a molecular weight of at the most 300 Da.

37. A method according to any of claims 20-36, wherein the substance is a peptide.

38. A method according to any of claim 20-36, wherein the substance is a carbon-containing non-peptide.

39. A method according to any of claims 20-38, wherein the organic compound is a compound having one or more chemical domains capable of interacting with one or more functional groups of the targeting sequence of the native anchoring site of the cyclic nucleotide phosphodiesterase.

40. A method according to claim 39, wherein the organic compound is a compound having at least two chemical domains capable of interacting with at least two functional groups of the targeting sequence of the native anchoring site for the cyclic nucleotide phosphodiesterase.

41. A method according to claim 39, wherein the organic compound is a compound having at least three chemical domains capable of interacting with at least three functional groups of the targeting sequence of the native anchoring site for the cyclic nucleotide phosphodiesterase.

42. A method according to any of claims 20-38, wherein the organic compound is a compound having one or more chemical domains capable of interacting with one or more functional groups of the targeting sequence of the cyclic nucleotide phosphodiesterase.

43. A method according to claim 42, wherein the organic compound is a compound having at least two chemical domains capable of interacting with at least two functional groups of the targeting sequence of the cyclic nucleotide phosphodiesterase.

44. A method according to claim 42, wherein the organic compound is a compound having at least three chemical domains capable of interacting with at least three functional groups of the targeting sequence of the cyclic nucleotide phosphodiesterase.

45. A method according to any of claims 20-44, wherein the organic compound is a weak acid in that it is a neutral molecule that can reversible dissociate into an anion (a negatively charged molecule) and a proton (a hydrogen ion).

46. A method according to claims 20-44, wherein the organic compound is a weak base in that it is a neutral molecule that can form a cation (a positively charged molecule) by combining with a proton (a hydrogen ion).

47. A method according to any of claims 39-46, wherein the functional groups of the targeting sequences include functional groups selected from the group consisting of: methyl-, isopropyl-, isobutyl-, hydroxyl-, thiol-, benzyl-, benzyloyl-, methyldolyl-, methylimidazolyl-, amine-, imine-, carboxyl- and acetamide-groups as parts of amino acids in the targeting sequences.

48. A method according to any of claims 20-47, wherein the exposure of the animal to the influence of a substance is performed by administering an effective amount of the substance to the animal.

49. A method according to claim 48, wherein the exposure of the animal to the influence of the substance is performed by administering an effective amount of the substance via the intravenous route of administration to the animal.

50. A method according to claim 48, wherein the exposure of the animal to the influence of the substance is performed by administering an effective amount of the substance via the oral route of administration to the animal.

51. A method according to claim 48, wherein the exposure of the animal to the influence of the substance is performed by administering an effective amount of the substance via the pulmonary route of administration to the animal.

52. A method according to claim 48, wherein the exposure of the animal to the influence of the substance is performed by administering an effective amount of the substance via the rectal route of administration to the animal.

53. A method according to claim 48, wherein the exposure of the animal to the influence of the substance is performed by administering an effective amount of the substance via the transdermal route of administration to the animal.

54. A method according to any of claims 21 - 53, wherein the quantitative fluorescence redistribution assay consists of cells selected from the group consisting of vascular smooth muscle cells such as vascular smooth muscle from resistance vessels on the arterial side of the vascular system, vascular smooth muscle from capacitance vessels on the venous side of the vascular system, vascular smooth muscle cells from small arteries, arterioles, venules, veins, any neuron from the brain, cardiomyocytes, vascular endothelial cells and immortal cells lines of these cells, and smooth vascular cells lines such as T/G HA-VSMCA10 and A7r5, transfected with a nucleotide construct encoding a fluorescent probe comprising as the biologically active polypeptide either splice variant mPDE 5, or PDE splice variant cloned from vascular smooth muscle cells such as vascular smooth muscle from resistance vessels on the arterial side of the vascular system, vascular smooth muscle from capacitance vessels on the venous side of the vascular system, vascular smooth muscle cells from small arteries, arterioles, venules, veins, any neuron from the brain, or cardiomyocytes, transfected in such a way, that the construct is expressed by the cells.

55. A method according to claim 54, wherein the quantitative fluorescence redistribution assay is a primary screening assay used in a discovery project

56. A method according to any of claim 54 or 55, wherein the cells are derived from an animal.

57. A method according to claim 56, wherein the cells are derived from a mammal such as a human.

58. A method according to any of claims 54-57, wherein the fluorescent probe redistributes after the cells have been subjected to a physiologically important and relevant influence that is relevant to the intercellular signalling pathway wherein the PDE splice variant is an integral part, so that both the normal pattern of spatial distribution and possible redistribution of the fluorescent probe can be established.

59. A method according to claim 58 wherein the intracellular signalling pathway comprises a cellular response that modulates the generation of cGMP and the spatially distributed cytoplasmic levels of cGMP.

60. A method according to any of claims 58 or 59, wherein the fluorescent probe is modified in a systematic way, still keeping the GFP coding sequence intact, so that the new fluorescent probes are fusion polypeptides where parts of the suspected targeting sequences of mPDE 5 are altered.

61. A method according to claim 60, wherein the modification of the suspected targeting sequence of the mPDE 5 or a splice variant is a deletion.

62. A method according to any of claims 60 or 61, wherein the spatial distribution of the fluorescent probe is compared to the spatial distribution of the unmodified fluorescent probe deducing the targeting sequence.

63. A method according to any of claims 20-62, wherein the substance interacts with the targeting sequence or part thereof in a manner that dislocates, disrupts targeting, or interferes with redistribution of the fluorescent probe as measured in quantitative fluorescence redistribution assay.

ABSTRACT

This application describes a method by which to identify novel chemical entities that may modulate the specific effectiveness of the cyclic nucleotide phosphodiesterases (PDEs) having the ability to cleave cGMP. The preferred mode of action is dislocation, disruption of targeting or interference with redistribution of specific isoforms or splice variants of PDE5 from their anchoring sites within cells, thereby modulating their specific effectiveness, not their enzymatic capacity. The chemical entities may be useful in preventing or treating, in an animal, preferably a human, in need thereof, an adverse condition which may be reduced or abolished by modulating the specific effectiveness of PDE5. Examples of such adverse conditions hypo- and hypertension, erectile dysfunction and jet-lag.

15 OKT. 1998

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: NovoNordisk, BioImage
- (ii) TITLE OF THE INVENTION: A method for preventing or treating adverse conditions which may be reduced or abolished by modulating the activity of one or more phosphodiesterases having the ability to cleave cGMP.
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: NovoNordisk, BioImage
 - (B) STREET: Mørkhøjbygade 28
 - (C) CITY: Søborg
 - (D) STATE: DK
 - (E) COUNTRY: DENMARK
 - (F) ZIP: 2860
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: , PV&P R
 - (B) REGISTRATION NUMBER:
 - (C) REFERENCE/DOCKET NUMBER:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3381 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
 - (A) NAME/KEY: Coding Sequence
 - (B) LOCATION: 1...3378
 - (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG GAG CGG GCC GGC CCC AGC TTC GGG CAG CAG CGA CAG CAG CAG

48

Met	Glu	Arg	Ala	Gly	Pro	Ser	Phe	Gly	Gln	Gln	Arg	Gln	Gln	Gln			
1				5				10					15				
CCC	CAG	CAG	CAG	AAG	CAG	CAG	CAG	AGG	GAT	CAG	GAC	TCG	GTC	GAA	GCA		96
Pro	Gln	Gln	Gln	Lys	Gln	Gln	Gln	Arg	Asp	Gln	Asp	Ser	Val	Glu	Ala		
			20					25					30				
TGG	CTG	GAC	GAT	CAC	TGG	GAC	TTT	ACC	TTC	TCA	TAC	TTT	GTT	AGA	AAA		144
Trp	Leu	Asp	Asp	His	Trp	Asp	Phe	Thr	Phe	Ser	Tyr	Phe	Val	Arg	Lys		
		35					40					45					
GCC	ACC	AGA	GAA	ATG	GTC	AAT	GCA	TGG	TTT	GCT	GAG	AGA	GTT	CAC	ACC		192
Ala	Thr	Arg	Glu	Met	Val	Asn	Ala	Trp	Phe	Ala	Glu	Arg	Val	His	Thr		
		50				55					60						
ATC	CCT	GTG	TGC	AAG	GAA	GGT	ATC	AGA	GGC	CAC	ACC	GAA	TCT	TGC	TCT		240
Ile	Pro	Val	Cys	Lys	Glu	Gly	Ile	Arg	Gly	His	Thr	Glu	Ser	Cys	Ser		
65				70					75					80			
TGT	CCC	TTG	CAG	CAG	AGT	CCT	CGT	GCA	GAT	AAC	AGT	GTC	CCT	GGA	ACA		288
Cys	Pro	Leu	Gln	Gln	Ser	Pro	Arg	Ala	Asp	Asn	Ser	Val	Pro	Gly	Thr		
			85					90					95				
CCA	ACC	AGG	AAA	ATC	TCT	GCC	TCT	GAA	TTT	GAC	CGG	CCT	CTT	AGA	CCC		336
Pro	Thr	Arg	Lys	Ile	Ser	Ala	Ser	Glu	Phe	Asp	Arg	Pro	Leu	Arg	Pro		
			100					105					110				
ATT	GTT	GTC	AAG	GAT	TCT	GAG	GGA	ACT	GTG	AGC	TTC	CTC	TCT	GAC	TCA		384
Ile	Val	Val	Lys	Asp	Ser	Glu	Gly	Thr	Val	Ser	Phe	Leu	Ser	Asp	Ser		
		115					120					125					
GAA	AAG	AAG	GAA	CAG	ATG	CCT	CTA	ACC	CCT	CCA	AGG	TTT	GAT	CAT	GAT		432
Glu	Lys	Lys	Glu	Gln	Met	Pro	Leu	Thr	Pro	Pro	Arg	Phe	Asp	His	Asp		
		130				135					140						
GAA	GGG	GAC	CAG	TGC	TCA	AGA	CTC	TTG	GAA	TTA	GTG	AAG	GAT	ATT	TCT		480
Glu	Gly	Asp	Gln	Cys	Ser	Arg	Leu	Leu	Glu	Leu	Val	Lys	Asp	Ile	Ser		
145				150					155					160			
AGT	CAT	TTG	GAT	GTC	ACA	GCC	TTA	TGT	CAC	AAA	ATT	TTC	TTG	CAT	ATC		528
Ser	His	Leu	Asp	Val	Thr	Ala	Leu	Cys	His	Lys	Ile	Phe	Leu	His	Ile		
			165					170					175				
CAT	GGA	CTG	ATA	TCT	GCT	GAC	CGC	TAT	TCC	CTG	TTC	CTT	GTC	TGT	GAA		576
His	Gly	Leu	Ile	Ser	Ala	Asp	Arg	Tyr	Ser	Leu	Phe	Leu	Val	Cys	Glu		
		180						185					190				
GAC	AGC	TCC	AAT	GAC	AAG	TTT	CTT	ATC	AGC	CGC	CTC	TTT	GAT	GTT	GCT		624
Asp	Ser	Ser	Asn	Asp	Lys	Phe	Leu	Ile	Ser	Arg	Leu	Phe	Asp	Val	Ala		
		195					200					205					
GAA	GGT	TCA	ACA	CTG	GAA	GAA	GTT	TCA	AAT	AAC	TGT	ATC	CGC	TTA	GAA		672
Glu	Gly	Ser	Thr	Leu	Glu	Glu	Val	Ser	Asn	Asn	Cys	Ile	Arg	Leu	Glu		
	210					215					220						

CAC CGT GGT GTG AAT AAC TCT TAC ATA CAG CGA AGT GAA CAT CCA CTT His Arg Gly Val Asn Asn Ser Tyr Ile Gln Arg Ser Glu His Pro Leu 660 665 670	2016
GCC CAG CTT TAC TGC CAT TCA ATC ATG GAA CAC CAT CAT TTT GAC CAG Ala Gln Leu Tyr Cys His Ser Ile Met Glu His His His Phe Asp Gln 675 680 685	2064
TGC CTG ATG ATT CTT AAT AGT CCA GGC AAT CAG ATT CTC AGT GGC CTC Cys Leu Met Ile Leu Asn Ser Pro Gly Asn Gln Ile Leu Ser Gly Leu 690 695 700	2112
TCC ATT GAA GAA TAT AAG ACC ACG TTG AAA ATA ATC AAG CAA GCT ATT Ser Ile Glu Glu Tyr Lys Thr Thr Leu Lys Ile Ile Lys Gln Ala Ile 705 710 715 720	2160
TTA GCT ACA GAC CTA GCA CTG TAC ATT AAG AGG CGA GGA GAA TTT TTT Leu Ala Thr Asp Leu Ala Leu Tyr Ile Lys Arg Arg Gly Glu Phe Phe 725 730 735	2208
GAA CTT ATA AGA AAA AAT CAA TTC AAT TTG GAA GAT CCT CAT CAA AAG Glu Leu Ile Arg Lys Asn Gln Phe Asn Leu Glu Asp Pro His Gln Lys 740 745 750	2256
GAG TTG TTT TTG GCA ATG CTG ATG ACA GCT TGT GAT CTT TCT GCA ATT Glu Leu Phe Leu Ala Met Leu Met Thr Ala Cys Asp Leu Ser Ala Ile 755 760 765	2304
ACA AAA CCC TGG CCT ATT CAA CAA CGG ATA GCA GAA CTT GTA GCA ACT Thr Lys Pro Trp Pro Ile Gln Gln Arg Ile Ala Glu Leu Val Ala Thr 770 775 780	2352
GAA TTT TTT GAT CAA GGA GAC AGA GAG AGA AAA GAA CTC AAC ATA GAA Glu Phe Phe Asp Gln Gly Asp Arg Glu Arg Lys Glu Leu Asn Ile Glu 785 790 795 800	2400
CCC ACT GAT CTA ATG AAC AGG GAG AAG AAA AAC AAA ATC CCA AGT ATG Pro Thr Asp Leu Met Asn Arg Glu Lys Lys Asn Lys Ile Pro Ser Met 805 810 815	2448
CAA GTT GGG TTC ATA GAT GCC ATC TGC TTG CAA CTG TAT GAG GCC CTG Gln Val Gly Phe Ile Asp Ala Ile Cys Leu Gln Leu Tyr Glu Ala Leu 820 825 830	2496
ACC CAC GTG TCA GAG GAC TGT TTC CCT TTG CTA GAT GGC TGC AGA AAG Thr His Val Ser Glu Asp Cys Phe Pro Leu Leu Asp Gly Cys Arg Lys 835 840 845	2544
AAC AGG CAG AAA TGG CAG GCC CTT GCA GAA CAG CAG GAG AAG ATG CTG Asn Arg Gln Lys Trp Gln Ala Leu Ala Glu Gln Gln Glu Lys Met Leu 850 855 860	2592
ATT AAT GGG GAA AGC GGC CAG GCC AAG CGG AAC TGG GTA CCG CGG GCC	2640

Ile Asn Gly Glu Ser Gly Gln Ala Lys Arg Asn Trp Val Pro Arg Ala	
865	870 875 880
CGG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC	2688
Arg Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe	
885 890 895	
ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC	2736
Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly	
900 905 910	
CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC	2784
His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly	
915 920 925	
AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC	2832
Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro	
930 935 940	
TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC	2880
Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser	
945 950 955 960	
CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG	2928
Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met	
965 970 975	
CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC	2976
Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly	
980 985 990	
AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG	3024
Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val	
995 1000 1005	
AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC	3072
Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile	
1010 1015 1020	
CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC	3120
Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile	
1025 1030 1035 1040	
ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC	3168
Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg	
1045 1050 1055	
CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG	3216
His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln	
1060 1065 1070	
AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC	3264
Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr	
1075 1080 1085	

CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT 3312
 Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp
 1090 1095 1100

CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC 3360
 His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly
 1105 1110 1115 1120

ATG GAC GAG CTG TAC AAG TAA 3381
 Met Asp Glu Leu Tyr Lys
 1125

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1126 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein
- (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Arg Ala Gly Pro Ser Phe Gly Gln Gln Arg Gln Gln Gln Gln
 1 5 10 15
 Pro Gln Gln Gln Lys Gln Gln Gln Arg Asp Gln Asp Ser Val Glu Ala
 20 25 30
 Trp Leu Asp Asp His Trp Asp Phe Thr Phe Ser Tyr Phe Val Arg Lys
 35 40 45
 Ala Thr Arg Glu Met Val Asn Ala Trp Phe Ala Glu Arg Val His Thr
 50 55 60
 Ile Pro Val Cys Lys Glu Gly Ile Arg Gly His Thr Glu Ser Cys Ser
 65 70 75 80
 Cys Pro Leu Gln Gln Ser Pro Arg Ala Asp Asn Ser Val Pro Gly Thr
 85 90 95
 Pro Thr Arg Lys Ile Ser Ala Ser Glu Phe Asp Arg Pro Leu Arg Pro
 100 105 110
 Ile Val Val Lys Asp Ser Glu Gly Thr Val Ser Phe Leu Ser Asp Ser
 115 120 125
 Glu Lys Lys Glu Gln Met Pro Leu Thr Pro Pro Arg Phe Asp His Asp
 130 135 140
 Glu Gly Asp Gln Cys Ser Arg Leu Leu Glu Leu Val Lys Asp Ile Ser
 145 150 155 160
 Ser His Leu Asp Val Thr Ala Leu Cys His Lys Ile Phe Leu His Ile
 165 170 175
 His Gly Leu Ile Ser Ala Asp Arg Tyr Ser Leu Phe Leu Val Cys Glu
 180 185 190
 Asp Ser Ser Asn Asp Lys Phe Leu Ile Ser Arg Leu Phe Asp Val Ala
 195 200 205
 Glu Gly Ser Thr Leu Glu Glu Val Ser Asn Asn Cys Ile Arg Leu Glu

210	215	220
Trp Asn Lys Gly Ile Val Gly His Val Ala Ala Leu Gly Glu Pro Leu		
225	230	235
Asn Ile Lys Asp Ala Tyr Glu Asp Pro Arg Phe Asn Ala Glu Val Asp		240
	245	250
Gln Ile Thr Gly Tyr Lys Thr Gln Ser Ile Leu Cys Met Pro Ile Lys		255
	260	265
Asn His Arg Glu Glu Val Val Gly Val Ala Gln Ala Ile Asn Lys Lys		270
	275	280
Ser Gly Asn Gly Gly Thr Phe Thr Glu Lys Asp Glu Lys Asp Phe Ala		285
	290	295
Ala Tyr Leu Ala Phe Cys Gly Ile Val Leu His Asn Ala Gln Leu Tyr		300
305	310	315
Glu Thr Ser Leu Leu Glu Asn Lys Arg Asn Gln Val Leu Leu Asp Leu		320
	325	330
Ala Ser Leu Ile Phe Glu Glu Gln Gln Ser Leu Glu Val Ile Leu Lys		335
	340	345
Lys Ile Ala Ala Thr Ile Ile Ser Phe Met Gln Val Gln Lys Cys Thr		350
	355	360
Ile Phe Ile Val Asp Glu Asp Cys Ser Asp Ser Phe Ser Ser Val Phe		365
	370	375
His Met Glu Cys Glu Glu Leu Glu Lys Ser Ser Asp Thr Leu Thr Arg		380
385	390	395
Glu His Asp Ala Asn Lys Ile Asn Tyr Met Tyr Ala Gln Tyr Val Lys		400
	405	410
Asn Thr Met Glu Pro Leu Asn Ile Pro Asp Val Ser Lys Asp Lys Arg		415
	420	425
Phe Pro Trp Thr Thr Glu Asn Thr Gly Asn Val Asn Gln Gln Cys Ile		430
	435	440
Arg Ser Leu Leu Cys Thr Pro Ile Lys Asn Gly Lys Lys Asn Lys Val		445
	450	455
Ile Gly Val Cys Gln Leu Val Asn Lys Met Glu Glu Asn Thr Gly Lys		460
465	470	475
Val Lys Pro Phe Asn Arg Asn Asp Glu Gln Phe Leu Glu Ala Phe Val		480
	485	490
Ile Phe Cys Gly Leu Gly Ile Gln Asn Thr Gln Met Tyr Glu Ala Val		495
	500	505
Glu Arg Ala Met Ala Lys Gln Met Val Thr Leu Glu Val Leu Ser Tyr		510
	515	520
His Ala Ser Ala Ala Glu Glu Glu Thr Arg Glu Leu Gln Ser Leu Ala		525
	530	535
Ala Ala Val Val Pro Ser Ala Gln Thr Leu Lys Ile Thr Asp Phe Ser		540
545	550	555
Phe Ser Asp Phe Glu Leu Ser Asp Leu Glu Thr Ala Leu Cys Thr Ile		560
	565	570
Arg Met Phe Thr Asp Leu Asn Leu Val Gln Asn Phe Gln Met Lys His		575
	580	585
Glu Val Leu Cys Arg Trp Ile Leu Ser Val Lys Lys Asn Tyr Arg Lys		590
	595	600
Asn Val Ala Tyr His Asn Trp Arg His Ala Phe Asn Thr Ala Gln Cys		605
610	615	620
Met Phe Ala Ala Leu Lys Ala Gly Lys Ile Gln Asn Lys Leu Thr Asp		625
	630	635
Leu Glu Ile Leu Ala Leu Leu Ile Ala Ala Leu Ser His Asp Leu Asp		640

49

His Arg Gly Val	645	Asn Asn Ser Tyr	650	Ile Gln Arg Ser Glu	655	His Pro Leu
	660		665		670	
Ala Gln Leu Tyr	675	Cys His Ser	680	Ile Met Glu His His	685	Phe Asp Gln
Cys Leu Met Ile	690	Leu Asn Ser	695	Pro Gly Asn Gln Ile	700	Leu Ser Gly Leu
Ser Ile Glu Glu	705	Tyr Lys Thr	710	Thr Leu Lys Ile	715	Ile Lys Gln Ala Ile
Leu Ala Thr Asp	725	Leu Ala Leu Tyr	730	Ile Lys Arg Arg	735	Gly Glu Phe Phe
Glu Leu Ile Arg	740	Lys Asn Gln Phe	745	Asn Leu Glu Asp	750	Pro His Gln Lys
Glu Leu Phe Leu	755	Ala Met Leu Met	760	Thr Ala Cys Asp	765	Leu Ser Ala Ile
Thr Lys Pro Trp	770	Pro Ile Gln Gln	775	Arg Ile Ala Glu	780	Leu Val Ala Thr
Glu Phe Phe Asp	785	Gln Gly Asp Arg	790	Glu Arg Lys Glu	795	Leu Asn Ile Glu
Pro Thr Asp Leu	805	Met Asn Arg Glu	810	Lys Lys Asn Lys	815	Ile Pro Ser Met
Gln Val Gly Phe	820	Ile Asp Ala Ile	825	Cys Leu Gln Leu	830	Tyr Glu Ala Leu
Thr His Val Ser	835	Glu Asp Cys Phe	840	Pro Leu Leu Asp	845	Gly Cys Arg Lys
Asn Arg Gln Lys	850	Trp Gln Ala Leu	855	Ala Glu Gln Gln	860	Glu Lys Met Leu
Ile Asn Gly Glu	865	Ser Gly Gln Ala	870	Lys Arg Asn Trp	875	Val Pro Arg Ala
Arg Asp Pro Pro	885	Val Ala Thr Met	890	Val Ser Lys Gly	895	Glu Glu Leu Phe
Thr Gly Val Val	900	Pro Ile Leu Val	905	Glu Leu Asp Gly	910	Asp Val Asn Gly
His Lys Phe Ser	915	Val Ser Gly Glu	920	Gly Glu Gly Asp	925	Ala Thr Tyr Gly
Lys Leu Thr Leu	930	Lys Phe Ile Cys	935	Thr Thr Gly Lys	940	Leu Pro Val Pro
Trp Pro Thr Leu	945	Val Thr Thr Leu	950	Thr Tyr Gly Val	955	Gln Cys Phe Ser
Arg Tyr Pro Asp	965	His Met Lys Gln	970	His Asp Phe Phe	975	Lys Ser Ala Met
Pro Glu Gly Tyr	980	Val Gln Glu Arg	985	Thr Ile Phe Phe	990	Lys Asp Asp Gly
Asn Tyr Lys Thr	995	Arg Ala Glu Val	1000	Lys Phe Glu Gly	1005	Asp Thr Leu Val
Asn Arg Ile Glu	1010	Leu Lys Gly Ile	1015	Asp Phe Lys Glu	1020	Asp Gly Asn Ile
Leu Gly His Lys	1025	Leu Glu Tyr Asn	1030	Tyr Asn Ser His	1035	Asn Val Tyr Ile
Met Ala Asp Lys	1045	Gln Lys Asn Gly	1050	Ile Lys Val Asn	1055	Phe Lys Ile Arg
His Asn Ile Glu	1060	Asp Gly Ser Val	1065	Gln Leu Ala Asp	1070	His Tyr Gln Gln
Asn Thr Pro Ile		Gly Asp Gly Pro		Val Leu Leu Pro		Asp Asn His Tyr

1075 1080 1085

Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp

1090 1095 1100

His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly

105 1110 1115 1120

Met Asp Glu Leu Tyr Lys

1125

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTGAATTCAA CCATGGAGCG GGCC

24

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTGGTACCCA GTTCGCTTG GCC

23

AN IMPROVED METHOD for extracting quantitative information relating to an influence on a cellular response.

SUMMARY OF THE INVENTION

5 The present invention relates to an improved method and tools for extracting quantitative information relating to an influence on a cellular response, in particular an influence caused by contacting or incubating the cell with a substance influencing a cellular response, wherein the cellular response is manifested in redistribution of at least one component in the cell. In particular, the invention relates to an improved method for
10 extracting the quantitative information relating to an influence on an intracellular pathway involving redistribution of at least one component associated with the pathway. The method of the invention may be used as a very efficient procedure for testing or discovering the influence of a substance on a physiological process, for example in connection with screening for new drugs, testing of substances for toxicity, identifying
15 drug targets for known or novel drugs. In particular, the present invention relates to an improved method for parallelisation of the testing procedure so that a large number of substances can be tested simultaneously using commercially available instrumentation. The invention also describes several ways of contacting the cells with a substance influencing a cellular response and modifications made to the actual cells before, during or
20 after contacting the cells with these substances as to improve the applicability and use of the method for extracting quantitative information relating to influence on an intracellular pathway in a highly parallel fashion. Other valuable uses of the method and technology of the invention will be apparent to the skilled person on the basis of the following disclosure. In a particular embodiment of the invention, the present invention relates to a method of
25 detecting intracellular translocation or redistribution of biologically active polypeptides, preferably an enzyme, affecting intracellular processes, and a DNA construct and a cell for use in the method.

Two appendices are included herein, and are considered part of the application. Appendix I, "METHOD AND APPARATUS FOR HIGH DENSITY FORMAT SCREENING FOR
30 BIOACTIVE MOLECULES", is a pending patent application. Appendix II, "CHANGES

IN INTRACELLULAR cAMP VISUALIZED USING A cAMP-DEPENDENT PROTEIN KINASE-GREEN FLUORESCENT PROTEIN HYBRID", is a manuscript intended for publication.

5 BACKGROUND OF THE INVENTION

Intracellular pathways are tightly regulated by a cascade of components that undergo modulation in a temporally and spatially characteristic manner. Several disease states can be attributed to altered activity of individual signalling components (i.e. protein kinases, protein phosphatases, transcription factors). These components therefore render
10 themselves as attractive targets for therapeutic intervention.

Protein kinases and phosphatases are well described components of several intracellular signalling pathways. The catalytic activity of protein kinases and phosphatases are assumed to play a role in virtually all regulatable cellular processes. Although the involvement of protein kinases in cellular signalling and regulation have been subjected to
15 extensive studies, detailed knowledge on e.g. the exact timing and spatial characteristics of signalling events is often difficult to obtain due to lack of a convenient technology.

Novel ways of monitoring specific modulation of intracellular pathways in intact, living cells is assumed to provide new opportunities in drug discovery, functional genomics, toxicology, patient monitoring etc.

20 The spatial orchestration of protein kinase activity is likely to be essential for the high degree of specificity of individual protein kinases. The phosphorylation mediated by protein kinases is balanced by phosphatase activity. Also within the family of phosphatases translocation has been observed, e.g. translocation of PTP2C to membrane ruffles [(Cossette *et al.* 1996)], and likewise is likely to be indicative of phosphatase activity.

25 Protein kinases often show a specific intracellular distribution before, during and after activation. Monitoring the translocation processes and/or redistribution of individual protein kinases or subunits thereof is thus likely to be indicative of their functional activity. A connection between translocation and catalytic activation has been shown for

protein kinases like the diacyl glycerol (DAG)-dependent protein kinase C (PKC), the cAMP-dependent protein kinase (PKA) [(DeBernardi *et al.* 1996)] and the mitogen-activated-protein kinase Erk-1 [(Sano *et al.* 1995)].

Commonly used methods of detection of intracellular localisation/activity of protein
5 kinases and phosphatases are immunoprecipitation, Western blotting and immunocytochemical detection.

Taking the family of diacyl glycerol (DAG)-dependent protein kinase Cs (PKCs) as an example, it has been shown that individual PKC isoforms that are distributed among different tissues and cells have different activator requirements and undergo differential
10 translocation in response to activation. Catalytically inactive DAG-dependent PKCs are generally distributed throughout the cytoplasm, whereas they upon activation translocate to become associated with different cellular components, e.g. plasma membrane [(Farese, 1992),(Fulop Jr. *et al.* 1995)] nucleus [(Khalil *et al.* 1992)], cytoskeleton [(Blobe *et al.* 1996)]. The translocation phenomenon being indicative of PKC activation has been
15 monitored using different approaches: a) immunocytochemistry where the localisation of individual isoforms can be detected after permeabilisation and fixation of the cells [(Khalil *et al.* 1992)]; and b) tagging all DAG-dependent PKC isoforms with a fluorescently labelled phorbol myristate acetate (PMA) [(Godson *et al.* 1996)]; and c) chemical tagging of PKC β 1 with the fluorophore Cy3 [(Bastiaens & Jovin 1996)] and d) genetic tagging of
20 PKC α [(Schmidt *et al.* 1997)] and of PKC γ and PKC δ [(Sakai *et al.* 1996)]. The first method does not provide dynamic information whereas the latter methods will. Tagging PKC with fluorescently labelled phorbol myristate acetate cannot distinguish between different DAG-dependent isoforms of PKC but will label and show movement of all isoforms. Chemical and genetic labelling of specific DAG-dependent PKCs confirmed that
25 they in an isoform specific manner upon activation move to cell periphery or nucleus.

In an alternative method, protein kinase A activity has been measured in living cells by chemical labelling one of the kinase's subunit [(Adams *et al.* 1991)]. The basis of the methodology is that the regulatory and catalytic subunit of purified protein kinase A is labelled with fluorescein and rhodamine, respectively. At low cAMP levels protein kinase
30 A is assembled in a heterotetrameric form which enables fluorescence resonance energy

transfer between the two fluorescent dyes. Activation of protein kinase A leads to dissociation of the complex, thereby eliminating the energy transfer. A disadvantage of this technology is that the labelled protein kinase A has to be microinjected into the cells of interest. This highly invasive technique is cumbersome and not applicable to large scale screening of biologically active substances. A further disadvantage of this technique as compared to the presented invention is that the labelled protein kinase A cannot be inserted into organisms/animals as a transgene.

Recently it was discovered that Green Fluorescent Protein (GFP) expressed in many different cell types, including mammalian cells, became highly fluorescent [(Chalfie *et al.* 1994)]. WO95/07463 describes a cell capable of expressing GFP and a method for detecting a protein of interest in a cell based on introducing into a cell a DNA molecule having DNA sequence encoding the protein of interest linked to DNA sequence encoding a GFP such that the protein produced by the DNA molecule will have the protein of interest fused to the GFP, then culturing the cells in conditions permitting expression of the fused protein and detecting the location of the fluorescence in the cell, thereby localizing the protein of interest in the cell. However, examples of such fused proteins are not provided, and the use of fusion proteins with GFP for detection or quantitation of translocation or redistribution of biologically active polypeptides affecting intracellular processes upon activation, such as proteins involved in signalling pathways, e.g. protein kinases or phosphatases, has not been suggested. WO 95/07463 further describes cells useful for the detection of molecules, such as hormones or heavy metals, in a biological sample, by operatively linking a regulatory element of the gene which is affected by the molecule of interest to a GFP, the presence of the molecules will affect the regulatory element which in turn will affect the expression of the GFP. In this way the gene encoding GFP is used as a reporter gene in a cell which is constructed for monitoring the presence of a specific molecular identity.

Green Fluorescent Protein has been used in an assay for the detection of translocation of the glucocorticoid receptor (GR) [(Carey, KL *et al.* 1996)]. A GR-S65TGFP fusion has been used to study the mechanisms involved in translocation of the glucocorticoid receptor (GR) in response to the agonist dexamethasone from the cytosol, where it is present in the absence of a ligand, through the nuclear pore to the nucleus where it remains after ligand

binding. The use of a GR-GFP fusion enables real-time imaging and quantitation of nuclear/cytoplasmic ratios of the fluorescence signal. A similar genetic construct has been used to follow and quantify dexamethasone induced translocation of GR to the nucleus in HeLa cells [(Guiliano, K.A *et al.* 1997)] in a system called Array Scan™ (WO 97/45730) designed for automated drug screening. Recently, several other investigators have demonstrated that tagging a specific protein (or part of a protein) involved in an intracellular signalling pathway with GFP provides a new means to measure and quantify the influence of substances on this pathway. The concept has been shown to work both for cytoplasmic to nuclear translocation of the androgen receptor [(Georget V *et al.* 1997)] and transcription factors such as NF-ATc [(Beals CR *et al.* 1997)] in analogy with what has already been described for GR above. Another relevant example is a β -arrestin – GFP construct that was shown to report on activation of G-protein coupled receptors by translocating from the cytosol to the plasma membrane [(Barak LS *et al.* 1997)]. Finally, it has also been demonstrated that attaching GFP to a smaller part of a protein like the pleckstrin homology domain of phospholipase C δ 1 [(Stauffer TP *et al.* 1998)] and a cysteine-rich domain of PKC γ [(Oancea E *et al.* 1998)] can be used to report on an influence from a substance by quantifying their redistribution within the cells during activation of the specific signalling pathway to which they belong.

Many currently used screening programmes designed to find compounds that affect protein kinase activity are based on measurements of kinase phosphorylation of artificial or natural substrates, receptor binding and/or reporter gene expression. The interest in fluorescence measurements as the basis for future high-throughput drug screening has however increased dramatically over the last few years [(Silverman L *et al.* 1998)]. Of particular interest to the present invention is a scanning laser imager for rapid screening of fluorescence changes in living cells [(Schroeder K & Neagle B 1996)] currently offered commercially by Molecular Devices, Inc. as the FLIPR™.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides an important new dimension in the investigation of cellular systems involving redistribution in that the invention provides quantification of the

redistribution responses or events caused by an influence, typically contact with a chemical substance or mixture of chemical substances, but also changes in the physical environment. The quantification makes it possible to set up meaningful relationships, expressed numerically, or as curves or graphs, between the influences (or the degree of influences) on cellular systems and the redistribution response. This is highly advantageous because, as has been found, the quantification can be achieved in both a fast and reproducible manner, and - what is perhaps even more important - the systems which become quantifiable utilising the method of the invention are systems from which enormous amounts of new information and insight can be derived.

10 The present screening assays have the distinct advantage over other screening assays, e.g., receptor binding assays, enzymatic assays, and reporter gene assays, in providing a system in which biologically active substances with completely novel modes of action, e.g. inhibition or promotion of redistribution/translocation of a biologically active polypeptide as a way of regulating its action rather than inhibition/activation of enzymatic activity, can be identified in a way that insures very high selectivity to the particular isoform of the biologically active polypeptide and further development of compound selectivity versus other isoforms of the same biologically active polypeptide or other components of the same signalling pathway.

In its broadest aspect, the invention relates to an improved method, with higher throughput compared to previous methods, for extracting quantitative information relating to an influence on a cellular response, the method comprising recording variation, caused by the influence on mechanically intact living cells, in spatially distributed light emitted from a luminophore, the luminophore being present in the cells and being capable of being redistributed in a manner which is related with the degree of the influence, and/or of being modulated by a component which is capable of being redistributed in a manner which is related to the degree of the influence, the association resulting in a modulation of the luminescence characteristics of the luminophore, detecting and recording the spatially distributed light from the luminophore, and processing the recorded variation in the spatially distributed light to provide quantitative information correlating the spatial distribution or change in the spatial distribution to the degree of the influence. In one aspect of the present invention the mechanically intact living cell is permeabilised at some

time after the influence has begun but during or before the actual experimental recording. In another aspect, the present invention relates to an improved method for extracting quantitative information relating to an influence on a cellular response, the method comprising recording variation, caused by the influence on permeabilised living cells, in spatially distributed light emitted from a luminophore, the luminophore being present in the cells and being capable of being redistributed in a manner which is related with the degree of the influence, and/or of being modulated by a component which is capable of being redistributed in a manner which is related to the degree of the influence, the association resulting in a modulation of the luminescence characteristics of the luminophore, detecting and recording the spatially distributed light from the luminophore, and processing the recorded variation in the spatially distributed light to provide quantitative information correlating the spatial distribution or change in the spatial distribution to the degree of the influence. In a preferred embodiment of the invention the luminophore, which is present in the cells, is capable of being redistributed by modulation of an intracellular pathway, in a manner which is related to the redistribution of at least one component of the intracellular pathway. In another preferred embodiment of the invention, the luminophore is a fluorophore.

In the invention the cell and/or cells are mechanically intact and alive throughout the experiment. In another embodiment of the invention, the cells are fixed at a point in time after the application of the influence at which the response has been predetermined to be significant, and the recording is made at an arbitrary later time. In another embodiment the cell and/or cells are mechanically intact and alive throughout the experiment but are mechanically or chemically disrupted or permeabilised as the initial step of experimental analysis. In another aspect of the invention the cells have their plasma membrane permanently and stably permeabilised before the initiation of the experiment in such a way that the plasma membrane stays permeable during the experiment. This allows the components of intracellular pathways to be contacted by substances that are not normally permeating the cell plasma membrane such as peptides, proteins and hydrophilic organic compounds.

The mechanically intact or permeabilised living cells could be selected from the group consisting of fungal cells, such as yeast cells; invertebrate cells including insect cells; and

vertebrate cells, such as mammalian cells. These cells are incubated at a temperature of 30°C or above, preferably at a temperature of from 32°C to 39°C, more preferably at a temperature of from 35°C to 38°C, and most preferably at a temperature of about 37°C during the time period over which the influence is observed. In one aspect of the invention the mechanically intact or permeabilised living cell is part of a matrix of identical or non-identical cells. In one embodiment of the invention the cells comprise a group or groups of cells contained within a spatial limitation or spatial limitations. In one embodiment, the cells comprise multiple groups of cells that are qualitatively the same but subjected to different influences. In another embodiment, the cells comprise multiple groups of cells that are qualitatively different but subjected to the same influence.

In one embodiment of the invention the spatial limitations are domains defined on a substrate on which the cells are present. The spatial limitations may be arranged in one or more arrays on a common carrier. The spatial limitations may be wells in a plate of microtiter type, such that 96, 384, 864 and 1536 wells are situated on the common carrier. In another embodiment the spatial limitations are wells in a plate of a format different from the microtiter type. In one embodiment of the invention the domains are established by the presence of the cells on the substrate in a pattern that defines the domains. In another aspect of the invention, the domains are instead established by the spatial pattern or array of the influence or influences as it/they are applied to or contacted by the cells. This aspect is thoroughly described in Appendix I. Briefly, in this aspect of the invention the mechanically intact or permeabilised living cells are part of a continuous or discontinuous sheet of cells cultured on an optically clear flat surface optimised or not for cell culture. The optically clear and flat surface may be a porous membrane that may allow cellular processes to grow through the membrane pores and may allow directed capillary flow of fluid through the pores.

A cell used in the present invention should contain a nucleic acid construct encoding a fusion polypeptide as defined herein and be capable of expressing the sequence encoded by the construct. The cell is a eukaryotic cell selected from the group consisting of fungal cells, such as yeast cells; invertebrate cells including insect cells; vertebrate cells such as mammalian cells. The preferred cells are mammalian cells.

In another aspect of the invention the cells could be from an organism carrying in at least one of its component cells a nucleic acid sequence encoding a fusion polypeptide as defined herein and be capable of expressing said nucleic acid sequence. The organism is selected from the group consisting of unicellular and multicellular organisms, such as a mammal.

The luminophore is the component that allows the redistribution to be visualised and/or recorded by emitting light in a spatial distribution related to the degree of influence. The term redistribution is intended to cover all aspects of a change in spatial location, such as a translocation of the luminophore or other components. In one embodiment of the invention, the luminophore is capable of being redistributed in a manner that is physiologically relevant to the degree of the influence. It should be understood that redistribution. In another embodiment, the luminophore is capable of associating with a component that is capable of being redistributed in a manner that is physiologically relevant to the degree of the influence. In another embodiment, a correlation between the redistribution of the luminophore and the degree of the influence could be determined experimentally. In a preferred aspect of the invention, the luminophore is capable of being redistributed in substantially the same manner as the at least one component of an intracellular pathway. In another embodiment of the invention, the luminophore is capable of being quenched upon spatial association with a component that is redistributed by modulation of the pathway, the quenching being measured as a change in the intensity of the luminescence. In another embodiment of the invention, the luminophore is stationary but may have a certain spatial distribution, and interacts with at least one component that is capable of being redistributed in a manner which is physiologically relevant to the degree of the influence, in such a way that one or more luminescence characteristics of the luminophore is/are modulated as the component moves closer to, or farther from, the luminophore.

The luminophore could be a fluorophore. In a preferred embodiment of the invention, the luminophore is a polypeptide encoded by and expressed from a nucleotide sequence harboured in the cells. The luminophore could be a hybrid polypeptide comprising a fusion of at least a portion of each of two polypeptides one of which comprises a luminescent polypeptide and the other one of which comprises a biologically active polypeptide, as

defined herein.

The luminescent polypeptide could be a GFP as defined herein or could be selected from the group consisting of green fluorescent proteins having the F64L mutation as defined herein such as F64L-GFP, F64L-Y66H-GFP, F64L-S65T-GFP, and EGFP. The GFP could
5 be N- or C-terminally tagged, optionally via a peptide linker, to the biologically active polypeptide or a part or a subunit thereof. The fluorescent probe could be a component of an intracellular signalling pathway. The probe is coded for by a nucleic acid construct.

The pathway of investigation in the present invention could be an intracellular signalling pathway.

10 In a preferred embodiment of the invention, the influence could be contact between the group or groups of mechanically intact or permeabilised living cells and a chemical substance, and/or incubation of the group or groups of mechanically intact or permeabilised living cells with a chemical substance in solution. In one aspect of the invention that is thoroughly described in Appendix I, the chemical substances are attached
15 to an underlying matrix. In this aspect, the chemical substances may also be produced and secreted from, or attached to the plasma membrane surfaces of, a sheet of genetically engineered cells. In this aspect of the invention the chemical substances may also have been separated two-dimensionally in a non-denaturing gel using electrophoresis and the gel is directly put in close proximity or direct contact with the mechanically intact or
20 permeabilised living cells so that the chemical substances can contact the cells through diffusion or convection.

The influence will modulate the intracellular processes. In one aspect the modulation could be an activation of the intracellular processes. In another aspect the modulation could be a deactivation of the intracellular processes. In yet another aspect, the influence could inhibit
25 or promote the redistribution without directly affecting the metabolic activity of the component of the intracellular processes.

In one embodiment the invention is used to establish a dose-response relationship for one or many chemical substances. In one embodiment the invention is used as a basis for a screening program, where the effect of unknown influences such as a compound library,

can be compared to influence of known reference compounds under standardised conditions.

In addition to the intensity, there are several parameters of fluorescence or luminescence that can be modulated by the effect of the influence on the underlying cellular phenomena, and can therefore be used in the invention. Some examples are resonance energy transfer, fluorescence lifetime, polarisation, and wavelength shift. Each of these methods requires a particular kind of filter in the emission light path to select the component of the light desired and reject other components. The recording of property of light could be in the form of an ordered array of values such as a CCD array or a vacuum tube device such as a vidicon. In addition, the translational mobility, or freedom of movement, of the luminophore attached to the protein of interest can be an important property affected by the influence on the underlying cellular phenomena, and can therefore be used in the invention.

In one embodiment of the invention, the spatially distributed light emitted by a luminophore is detected by a change in the resonance energy transfer between the luminophore and another luminescent entity capable of delivering energy to the luminophore, each of which has been selected or engineered to become part of, bound to or associated with particular components of the intracellular pathway. In this embodiment, either the luminophore or the luminescent entity capable of delivering energy to the luminophore undergoes redistribution in response to an influence. The resonance energy transfer would be measured as a change in the intensity of emission from the luminophore, preferably sensed by a single channel photodetector that responds only to the average intensity of the luminophore in a non-spatially resolved fashion.

In one embodiment of the invention, the spatially distributed light emitted by a luminophore includes the case of uniform spatial distribution of the light.

In one aspect of the invention, the luminophore is a fluorophore which redistributes through a non-homogenous excitation light field, resulting in a change in the intensity of the light emitted from the luminophore as a result of the change in the amount of excitation light intensity at different points in the field.

In one embodiment of the invention, the recording of the spatially distributed light could be made at a single point in time after the application of the influence. In another embodiment, the recording could be made at two points in time, one point being before, and the other point being after the application of the influence. The result or variation is determined from the change in fluorescence compared to the fluorescence measured prior to the influence or modulation. In another embodiment of the invention, the recording could be performed at a series of points in time, in which the application of the influence occurs at some time after the first time point in the series of recordings, the recording being performed, e.g., with a predetermined time spacing of from 0.1 seconds to 1 hour, preferably from 1 to 60 seconds, more preferably from 1 to 30 seconds, in particular from 1 to 10 seconds, over a time span of from 1 second to 12 hours, such as from 10 seconds to 12 hours, e.g., from 10 seconds to one hour, such as from 60 seconds to 30 minutes or 20 minutes. The result or variation is determined from the change in fluorescence over time. The result or variation could also be determined as a change in the spatial distribution of the fluorescence over time.

In one embodiment the recording comprises a time series of total luminescence of the cells of one or several of the spatial limitations. In one embodiment the signal from all of the spatial limitations, one at a time, is measured by a recording being made in the individual spatial limitations by means of an apparatus to sequentially position each one of the limitations in the field of view of the detector and repeating the positioning and measurement process until all of the spatial limitations have been measured. The detector may be a photomultiplier tube. In a preferred embodiment of the present invention more than one spatial limitation is measured simultaneously. This may be done by means of a one- or two-dimensional array detector, whereby the multiple spatial limitations are imaged onto the array detector such that discrete subsets of the detecting units (pixels) in the array detector measure the signal from one and only one of the multiple spatial limitations, the signal from any one spatial limitation being the combined signal from those pixels that receive the image from one of the spatial limitations. This array detector may be a linear diode array, a video camera (according to any present or future standards and definitions of image acquisition and transmission) or a charge transfer device such as a charge-coupled device (CCD). In one embodiment the recording of signal requires

illumination of the multiple spatial limitations to excite the luminophores so that they emit light. In one embodiment all of the spatial limitations are simultaneously illuminated during the measurement. In another embodiment the spatial limitations are singly illuminated only during the time in which they are being measured. In a preferred embodiment the illumination is provided by a laser that is scanned in a raster fashion over some or all of the spatial limitations being measured. The scanning may take place at a rate that is substantially faster than the measurement process such that the illumination appears to the measurement process to be continuous in time and spatially uniform over the region being measured.

- 10 The recording of spatially distributed luminescence emitted from the luminophore is performed by an apparatus for measuring the distribution of fluorescence in the cells, and thereby any change in the distribution of fluorescence in the cells, which includes at a minimum the following component parts: (a) a light source, (b) a method for selecting the wavelength(s) of light from the source which will excite the luminescence of the
- 15 luminophore, (c) a device which can rapidly block or pass the excitation light into the rest of the system, (d) a series of optical elements for conveying the excitation light to the specimen, collecting the emitted fluorescence in a spatially resolved fashion, and forming an image from this fluorescence emission (or another type of intensity map relevant to the method of detection and measurement), (e) a bench or stand which holds the container of
- 20 the cells being measured in a predetermined geometry with respect to the series of optical elements, (f) a detector to record the spatially resolved fluorescence in the form of an image, (g) a computer or electronic system and associated software to acquire and store the recorded images, and to compute the degree of redistribution from the recorded images.

In a preferred embodiment of the invention the apparatus system is automated. In one embodiment the components in d and e mentioned above comprise a fluorescence microscope. In one embodiment the component in f mentioned above is a CCD camera. In one embodiment the component in f mentioned above is an array of photomultiplier tubes/devices.

In one embodiment the image is formed and recorded by an optical scanning system.

- 30 In one embodiment the optical scanning system is used to illuminate the bottom of a plate

of microtiter type so that a time-resolved recording of changes in luminescence or fluorescence can be made from all spatial limitations simultaneously.

In a preferred embodiment the actual luminescence or fluorescence measurements are made in a FLIPR™ instrument, commercially available from Molecular Devices, Inc.

- 5 In one embodiment of the invention the actual fluorescence measurements are made in a standard type of fluorometer for plates of microtiter type (fluorescence plate reader).

In one embodiment a liquid addition system is used to add a known or unknown compound to any or all of the cells in the cell holder at a time determined in advance. Preferably, the liquid addition system is under the control of the computer or electronic system. Such an
10 automated system can be used for a screening program due to its ability to generate results from a larger number of test compounds than a human operator could generate using the apparatus in a manual fashion.

The methods whereby the detector layer of cells are physically contacted by the compounds can also be of another conceptual type where the compounds are delivered to
15 the cells through a porous membrane by convection/diffusion or by directly contacting compounds attached to an inorganic or organic support (such as glass, plastic or the plasma membrane of intact living cells) with the cells. These methods are thoroughly described in Appendix I, but are also outlined in the following paragraphs.

In one aspect of the present invention where the detector layer of cells is a continuous or
20 discontinuous sheet of cells without any separation into test units or wells. The compounds are printed onto a nonabsorbent sheet of porous material as a solution in solvent and allowed to dry. This printed sheet of compounds then defines the test pattern for the experiment as it is brought down in close proximity to or in direct contact with the underlying detector layer of cells. The compounds, now dissolved by the fluid layer on the
25 cells, is brought in contact with the cells through the pores of the membrane by convection. The porous membrane onto which the compounds are printed is optically clear and preferably composed as stated in Appendix I. In another embodiment of this aspect of the present invention the detector layer of cells is a continuous or discontinuous sheet of cells, without any separation into test units or wells, growing on a porous and optically clear

membrane preferably of the types mentioned above. The porous membrane may allow the cells to send cellular processes through the pores of the membrane. The compounds are printed onto an optically clear substratum such as glass, plastic or quartz as solutions in solvent and allowed to dry. At the time of the experiment the cell sheet on the membrane, surrounded by a thin film of fluid, is layered on top of the printed compound pattern. The compounds then dissolve and contact the cells via diffusion and convection. The compounds may be made using combinatorial chemistry techniques, and may be peptides. The compounds may be covalently attached to the optically clear substratum or porous membrane. The compounds may also be proteins, polypeptides or peptides secreted by or attached to the plasma membrane of genetically modified cells growing as a continuous or discontinuous sheet on a flat optically clear surface or an optically clear porous membrane.

The recording of the variation or result with respect to light emitted from the luminophore is performed by recording the spatially distributed light as one or more digital images, and the processing of the recorded variation to reduce it to one or more numbers representative of the degree of redistribution comprises a digital image processing procedure or combination of digital image processing procedures. The quantitative information which is indicative of the degree of the cellular response to the influence or the result of the influence on the intracellular pathway is extracted from the recording or recordings according to a predetermined calibration based on responses or results, recorded in the same manner, to known degrees of a relevant specific influence. This calibration procedure is developed according to principles described below (Developing an Image-based Assay Technique). Specific descriptions of the procedures for particular assays are given in the examples.

While the stepwise procedure necessary to reduce the image or images to the value representative of the response caused by the influence is particular to each assay, the individual steps are generally well-known methods of image processing. Some examples of the individual steps are point operations such as subtraction, ratioing, and thresholding, digital filtering methods such as smoothing, sharpening, and edge detection, spatial frequency methods such as Fourier filtering, image cross-correlation and image autocorrelation, object finding and classification (blob analysis), and colour space manipulations for visualisation. In addition to the algorithmic procedures, heuristic

methods such as neural networks may also be used. In a preferred embodiment of the invention, a dose-response relationship is established based on quantification of the responses caused by a particular influence, representative of the underlying intracellular signalling process, using the methods described above and in examples 1-22 and 25. The dose-response relationship for the particular influence is then compared to the dose-response relationship obtained by performing the same assay in an instrument which allows parallel monitoring of all wells in a microtiter plate such as a FLIPR™ or an ordinary fluorescence plate reader for microtiter plates. If a good correlation between the dose-response relationships obtained from the two different measurement systems is obtained, it can be said that the parallel measurement mode has been validated (see examples 23 and 24). This implies that it can be used as the primary basis for a screening assay with the potential benefit of screening a significantly higher number of substances per unit of time for their influence on the response.

Imaging plate readers integrate the signal from each well into a single value per time point. Thus the data resulting from a single “run” of the instrument is a set of time series of single values, one for each well, with the injection of the test compound taking place at a known point in the time series. The primary advantage of this type of instrumentation is that it greatly increases the number of samples that can be processed in a given amount of time (the throughput). This is of great advantage when using the assay in a screening program for new pharmaceutical lead compounds.

The first step in the data analysis is to normalise the results from each well so that they can be compared with each other or with previously analysed known compounds. This always begins with correcting the signal by subtracting the instrument bias from all data points on a well-by-well basis. From this point, either of two techniques can be followed depending on the design of the assay:

Procedure 1: The average of the signal prior to the addition of the test compound is subtracted from all data points on a well-by-well basis.

Procedure 2: The data are corrected for any known background by subtracting the background value from all data points on a well-by-well basis. The resulting background-corrected data are normalised by dividing each data set by the average of the data values

prior to the injection of the test compound on a well-by-well basis.

The corrected or normalised time series data sets are then further reduced by a technique that converts the time series to a single value. There are at least three such approaches:

1. For transient responses, the maximum deviation from the baseline is determined. This is also known as the "peak height" technique.
2. Alternatively, the signal is integrated over time between pre-defined limits. If the data were treated according to Procedure 2 above, then the offset is subtracted such that the integral of a non-response is zero within the limit of measurement error. This is also known as the "peak area" technique.
3. If the response is a cumulative one, e.g., an exponential change to a new level, the result is taken as either the difference or the ratio between the signal after a predetermined time and the signal prior to the addition of the test compound.

All of the above procedures reduce the data for a given well to one or more single values. For screening purposes, these values will be searched for those that are greater than a certain statistically determined cut-off value. For characterisation, the values represent a quantitative response, and are further treated in sets by techniques such as dose-response curve fitting.

In another embodiment of the invention, the measurement of redistribution is accomplished indirectly by taking advantage of the fact that in order for redistribution to occur, the probe will experience some change in its freedom, or restriction, of movement within the intracellular milieu. The degree of translocation will correlate with the amount of freely mobile luminophore in the cytoplasm. At a point in time after the test compound has begun to have any influence it may have, the amount or fraction of restricted luminophore can be measured by disrupting or permeabilising the plasma membrane of the cells and allowing the freely mobile luminophore to diffuse away. If the detection volume of the detector is limited to the region immediately surrounding the cells, and the overall volume into which the freely mobile luminophore can diffuse is much larger, then the freely mobile luminophore essentially disappears from the detector's view and its signal is

not recorded.

In one aspect of the invention, the above mentioned measurement of redistribution is made on cells with permanently permeabilised plasma membranes immersed in a solution mimicking the cytoplasmic environment. In this way the influence of compounds that can normally not enter the cytoplasm of cells can be tested.

The nucleic acid constructs used in the present invention encode in their nucleic acid sequences fusion polypeptides comprising a biologically active polypeptide that is a component of an intracellular signalling pathway, or a part thereof, and a GFP, preferably an F64L mutant of GFP, N- or C-terminally fused, optionally via a peptide linker, to the biologically active polypeptide or part thereof.

In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a protein kinase or a phosphatase.

In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a transcription factor or a part thereof which changes cellular localisation upon activation.

In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a protein, or a part thereof, which is associated with the cytoskeletal network and which changes cellular localisation upon activation.

In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a protein kinase or a part thereof which changes cellular localisation upon activation.

In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a serine/threonine protein kinase or a part thereof capable of changing intracellular localisation upon activation.

In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a tyrosine protein kinase or a part thereof capable of changing intracellular localisation upon activation.

In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a phospholipid-dependent serine/threonine protein kinase or a part thereof capable of changing intracellular localisation upon activation.

5 In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a cAMP-dependent protein kinase or a part thereof capable of changing cellular localisation upon activation. In a preferred embodiment the biologically active polypeptide encoded by the nucleic acid construct is a PKAc-F64L-S65T-GFP fusion.

10 In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a cGMP-dependent protein kinase or a part thereof capable of changing cellular localisation upon activation.

In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a calmodulin-dependent serine/threonine protein kinase or a part thereof capable of changing cellular localisation upon activation.

15 In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a mitogen-activated serine/threonine protein kinase or a part thereof capable of changing cellular localisation upon activation. In preferred embodiments the biologically active polypeptide encoded by the nucleic acid constructs are an ERK1-F64L-S65T-GFP fusion or an EGFP-ERK1 fusion.

20 In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a cyclin-dependent serine/threonine protein kinase or a part thereof capable of changing cellular localisation upon activation.

In one embodiment the biologically active polypeptide encoded by the nucleic acid construct is a protein phosphatase or a part thereof capable of changing cellular localisation upon activation.

25 In one preferred embodiment of the invention the nucleic acid constructs may be DNA constructs.

In one embodiment the biologically active polypeptide encoded by the nucleic acid

construct. In one embodiment the gene encoding GFP in the nucleic acid construct is derived from *Aequorea victoria*. In a preferred embodiment the gene encoding GFP in the nucleic acid construct is EGFP or a GFP variant selected from F64L-GFP, F64L-Y66H-GFP and F64L-S65T-GFP.

- 5 In preferred embodiments of the invention the DNA constructs which can be identified by any of the DNA sequences shown in SEQ ID NO: 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, and 152 or are variants of these sequences capable of encoding the same fusion polypeptide or a fusion polypeptide which
- 10 is biologically equivalent thereto, e.g. an isoform, or a splice variant or a homologue from another species.

The present invention describes a method that may be used to establish a screening program for the identification of biologically active substances that directly or indirectly affects intracellular signalling pathways and because of this property are potentially useful

15 as medicaments. Based on measurements in living cells of the redistribution of spatially resolved luminescence from luminophores which undergo a change in distribution upon activation or deactivation of an intracellular signalling pathway the result of the individual measurement of each substance being screened indicates its potential biological activity.

- In one embodiment of the invention the screening program is used for the identification of
- 20 a biologically toxic substance as defined herein that exerts its toxic effect by interfering with an intracellular signalling pathway. Based on measurements in living cells of the redistribution of spatially resolved luminescence from luminophores which undergo a change in distribution upon activation or deactivation of an intracellular signalling pathway the result of the individual measurement of each substance being screened
- 25 indicates its potential biologically toxic activity. In one embodiment of a screening program a compound that modulates a component of an intracellular pathway as defined herein, can be found and the therapeutic amount of the compound estimated by a method according to the method of the invention. In a preferred embodiment the present invention leads to the discovery of a new way of treating a condition or disease related to the
- 30 intracellular function of a biologically active polypeptide comprising administration to a

patient suffering from said condition or disease of an effective amount of a compound which has been discovered by any method according to the invention. In another preferred embodiment of the invention a method is established for identification of a new drug target or several new drug targets among the group of biologically active polypeptides which are components of intracellular signalling pathways.

In another embodiment of the invention an individual treatment regimen is established for the selective treatment of a selected patient suffering from an ailment where the available medicaments used for treatment of the ailment are tested on a relevant primary cell or cells obtained from said patient from one or several tissues, using a method comprising

transfecting the cell or cells with at least one DNA sequence encoding a fluorescent probe according to the invention, transferring the transfected cell or cells back the said patient, or culturing the cell or cells under conditions permitting the expression of said probes and exposing it to an array of the available medicaments, then comparing changes in fluorescence patterns or redistribution patterns of the fluorescent probes in the intact living cells to detect the cellular response to the specific medicaments (obtaining a cellular action profile), then selecting one or more medicament or medicaments based on the desired activity and acceptable level of side effects and administering an effective amount of these medicaments to the selected patient.

The present invention describes a method that may be used to establish a screening program for back-tracking signal transduction pathways as defined herein. In one embodiment the screening program is used to establish more precisely at which level one or several compounds affect a specific signal transduction pathway by successively or in parallel testing the influence of the compound or compounds on the redistribution of spatially resolved luminescence from several of the luminophores which undergo a change in distribution upon activation or deactivation of the intracellular signalling pathway under study.

In general, a probe, i.e. a "GeneX"-GFP fusion or a GFP-"GeneX" fusion, is constructed using PCR with "GeneX"-specific primers followed by a cloning step to fuse "GeneX" in frame with GFP. The fusion may contain a short vector derived sequence between "GeneX" and GFP (e.g. part of a multiple cloning site region in the plasmid) resulting in a

peptide linker between "GeneX" and GFP in the resulting fusion protein.

Some of the steps involved in the development of a probe include the following:

- Identify the sequence of the gene. This is most readily done by searching a depository of genetic information, e.g. the GenBank Sequence Database, which is widely available and routinely used by molecular biologists. In the specific examples below the GenBank Accession number of the gene in question is provided.
- Design the gene-specific primers. Inspection of the sequence of the gene allows design of gene-specific primers to be used in a PCR reaction. Typically, the top-strand primer encompasses the ATG start codon of the gene and the following ca. 20 nucleotides, while the bottom-strand primer encompasses the stop codon and the ca. 20 preceding nucleotides, if the gene is to be fused behind GFP, i.e. a GFP-"GeneX" fusion. If the gene is to be fused in front of GFP, i.e. a "GeneX"-GFP fusion, a stop codon must be avoided. Optionally, the full-length sequence of GeneX may not be used in the fusion, but merely the part that localizes and redistributes like GeneX in response to a signal. In addition to gene-specific sequences, the primers contain at least one recognition sequence for a restriction enzyme, to allow subsequent cloning of the PCR product. The sites are chosen so that they are unique in the PCR product and compatible with sites in the cloning vector. Furthermore, it may be necessary to include an exact number of nucleotides between the restriction enzyme site and the gene-specific sequence in order to establish the correct reading frame of the fusion gene and/or a translation initiation consensus sequence. Lastly, the primers always contain a few nucleotides in front of the restriction enzyme site to allow efficient digestion with the enzyme.
- Identify a source of the gene to be amplified. In order for a PCR reaction to produce a product with gene-specific primers, the gene-sequence must initially be present in the reaction, e.g. in the form of cDNA. Information in GenBank or the scientific literature will usually indicate in which tissue(s) the gene is expressed, and cDNA libraries from a great variety of tissues or cell types from various species are commercially available, e.g. from Clontech (Palo Alto), Stratagene (La Jolla) and Invitrogen (San Diego). Many genes are also available in cloned form from The American Type Tissue

Collection (Virginia).

- Optimise the PCR reaction. Several factors are known to influence the efficiency and specificity of a PCR reaction, including the annealing temperature of the primers, the concentration of ions, notably Mg^{2+} and K^+ , present in the reaction, as well as pH of the reaction. If the result of a PCR reaction is deemed unsatisfactory, it might be because the parameters mentioned above are not optimal. Various annealing temperatures should be tested, e.g. in a PCR machine with a built-in temperature gradient, available from e.g. Stratagene (La Jolla), and/or various buffer compositions should be tried, e.g. the OptiPrime buffer system from Stratagene (La Jolla).
- Clone the PCR product. The vector into which the amplified gene product will be cloned and fused with GFP will already have been taken into consideration when the primers were designed. When choosing a vector, one should at least consider in which cell types the probe subsequently will be expressed, so that the promoter controlling expression of the probe is compatible with the cells. Most expression vectors also contain one or more selective markers, e.g. conferring resistance to a drug, which is a useful feature when one wants to make stable transfectants. The selective marker should also be compatible with the cells to be used.

The actual cloning of the PCR product should present no difficulty as it typically will be a one-step cloning of a fragment digested with two different restriction enzymes into a vector digested with the same two enzymes. If the cloning proves to be problematic, it may be because the restriction enzymes did not work well with the PCR fragment. In this case one could add longer extensions to the end of the primers to overcome a possible difficulty of digestion close to a fragment end, or one could introduce an intermediate cloning step not based on restriction enzyme digestion. Several companies offer systems for this approach, e.g. Invitrogen (San Diego) and Clontech (Palo Alto).

Once the gene has been cloned and, in the process, fused with the GFP gene, the resulting product, usually a plasmid, should be carefully checked to make sure it is as expected. The most exact test would be to obtain the nucleotide sequence of the fusion-gene.

Once a DNA construct for a probe has been generated, its functionality and usefulness may

be evaluated by transfecting it into cells capable of expressing the probe. The fluorescence of the cell is inspected soon after, typically the next day. At this point, two features of cellular fluorescence are noted: the intensity and the sub-cellular localisation.

The intensity should usually be at least as strong as that of unfused GFP in the cells. If it is not, the sequence or quality of the probe-DNA might be faulty, and should be carefully checked.

The sub-cellular localisation is an indication of whether the probe is likely to perform well. If it localises as expected for the gene in question, e.g. is excluded from the nucleus, it can immediately go on to a functional test. If the probe is not localised soon after the transfection procedure, it may be because of overexpression at this point in time, as the cell typically will have taken up very many copies of the plasmid, and localisation will occur in time, e.g. within a few weeks, as plasmid copy number and expression level decreases. If localisation does not occur after prolonged time, it may be because the fusion to GFP has destroyed a localisation function, e.g. masked a protein sequence essential for interaction with its normal cellular anchor-protein. In this case the opposite fusion might work, e.g. if GeneX-GFP does not work, GFP-GeneX might, as two different parts of GeneX will be affected by the proximity to GFP. If this does not work, the proximity of GFP at either end might be a problem, and it could be attempted to increase the distance by incorporating a longer linker between GeneX and GFP in the DNA construct.

If there is no prior knowledge of localisation, and no localisation is observed, it may be because the probe should not be localised at this point, because such is the nature of the protein fused to GFP. It should then be subjected to a functional test.

In a functional test, the cells expressing the probe are treated with at least one compound known to perturb, usually by activating, the signalling pathway on which the probe is expected to report by redistributing itself within the cell. If the redistribution is as expected, e.g. if prior knowledge tell that it should translocate from location X to location Y, it has passed the first critical test. In this case it can go on to further characterisation and quantification of the response.

If it does not perform as expected, it may be because the cell lacks at least one component

of the signalling pathway, e.g. a cell surface receptor, or there is species incompatibility, e.g. if the probe is modelled on sequence information of a human gene product, and the cell is of hamster origin. In both instances one should identify other cell types for the testing process where these potential problems would not apply.

- 5 If there is no prior knowledge about the pattern of redistribution, the analysis of the redistribution will have to be done in greater depth to identify what the essential and indicative features are, and when this is clear, it can go on to further characterisation and quantification of the response. If no feature of redistribution can be identified, the problem might be as mentioned above, and the probe should be retested under more optimal cellular
- 10 conditions.

If the probe does not perform under optimal cellular conditions, then it's back to the drawing board.

- The process of developing an image-based redistribution assay begins with either the unplanned experimental observation that a redistribution phenomenon can be visualised, or
- 15 the design of a probe specifically to follow a redistribution phenomenon already known to occur. In either event, the first and best exploratory technique is for a trained scientist or technician to observe the phenomenon. Even with the rapid advances in computing technology, the human eye-brain combination is still the most powerful pattern recognition system known, and requires no advance knowledge of the system in order to detect
- 20 potentially interesting and useful patterns in raw data. This is especially if those data are presented in the form of images, which are the natural "data type" for human visual processing. Because human visual processing operates most effectively in a relatively narrow frequency range, i.e., we cannot see either very fast or very slow changes in our visual field, it may be necessary to record the data and play it back with either time
- 25 dilation or time compression.

- Some luminescence phenomena cannot be seen directly by the human eye. Examples include polarisation and fluorescence lifetime. However, with suitable filters or detectors, these signals can be recorded as images or sequences of images and displayed to the human in the fashion just described. In this way, patterns can be detected and the same
- 30 methods can be applied.

Once the redistribution has been determined to be a reproducible phenomenon, one or more data sets are generated for the purpose of developing a procedure for extracting the quantitative information from the data. In parallel, the biological and optical conditions are determined which will give the best quality raw data for the assay. This can become an
5 iterative process; it may be necessary to develop a quantitative procedure in order to assess the effect on the assay of manipulating the assay conditions.

The data sets are examined by a person or persons with knowledge of the biological phenomenon and skill in the application of image processing techniques. The goal of this exercise is to determine or at least propose a method that will reduce the image or
10 sequence of images constituting the record of a "response" to a value corresponding to the degree of the response. Using either interactive image processing software or an image processing toolbox and a programming language, the method is encoded as a procedure or algorithm that takes the image or images as input and generates the degree of response (in any units) as its output. Some of the criteria for evaluating the validity of a particular
15 procedure are:

- Does the degree of the response vary in a biologically significant fashion, i.e., does it show the known or putative dependence on the concentration of the stimulating agent or condition?
- Is the degree of response reproducible, i.e., does the same concentration or level of
20 stimulating agent or condition give the same response with an acceptable variance?
- Is the dynamic range of the response sufficient for the purpose of the assay? If not, can a change in the procedure or one of its parameters improve the dynamic range?
- Does the procedure exhibit any clear "pathologies", i.e., does it give ridiculous values for the response if there are commonly occurring imperfections in the
25 imaging process? Can these pathologies be eliminated, controlled, or accounted for?
- Can the procedure deal with the normal variation in the number and/or size of cells in an image?

In some cases the method may be obvious; in others, a number of possible procedures may suggest themselves. Even if one method appears clearly superior to others, optimisation of parameters may be required. The various procedures are applied to the data set and the criteria suggested above are determined, or the single procedure is applied repeatedly with
5 adjustment of the parameter or parameters until the most satisfactory combination of signal, noise, range, etc. are arrived at. This is equivalent to the calibration of any type of single-channel sensor.

The number of ways of extracting a single value from an image are extremely large, and thus an intelligent approach must be taken to the initial step of reducing this number to a
10 small, finite number of possible procedures. This is not to say that the procedure arrived at is necessarily the best procedure - but a global search for the best procedure is simply out of the question due to the sheer number of possibilities involved.

Image-based assays are no different than other assay techniques in that their usefulness is characterised by parameters such as the specificity for the desired component of the
15 sample, the dynamic range, the variance, the sensitivity, the concentration range over which the assay will work, and other such parameters. While it is not necessary to characterise each and every one of these before using the assay, they represent the only way to compare one assay with another.

The final step is then to see whether there exists a possibility to increase the throughput of
20 the assay to improve its utility as the basis of a screening program. In order to do this, a dose-response relationship is established based on quantification of the responses caused by a particular influence, representative of the underlying intracellular signalling process, using the methods described above and in examples 1-22 and 25. The dose-response
25 relationship for the particular influence is then compared to the dose-response relationship obtained by performing the same assay in an instrument which allows parallel monitoring of all wells in a microtiter plate such as a FLIPR™ or an ordinary imaging or fluorescence plate reader for microtiter plates. If a good correlation between the dose-response
relationships obtained from the two different measurement systems is obtained, it can be
30 said that the parallel measurement mode has been validated (see examples 23 and 24). This implies that it can be used as the primary basis for a screening program with the potential

benefit of screening a significantly higher number of substances for their influence on the response per unit of time.

The process of developing an image-based assay is best illustrated by example. The development of such an assay for GLUT4 translocation is hereby described. GLUT4 is a member of the class of glucose transporter molecules that are important in cellular glucose uptake. It is known to translocate to the plasma membrane under some conditions of stimulation of glucose uptake. The ability to visualise the glucose uptake response non-invasively, without actually measuring glucose uptake, would be a very useful assay for anyone looking for, for example, treatments for type II diabetes.

- 10 A CHO cell line which stably expressed the human insulin receptor was used as the basis for a new cell line which stably expressed a fusion between GLUT4 and GFP. This cell line was expected to show translocation of GLUT4 to the plasma membrane as visualised by the movement of the GFP. The translocation could definitely be seen in the form of the appearance of local increases in the fluorescence in regions of the plasma membrane which had a characteristic shape or pattern. This is shown in Figure 12.

- These objects became known as "snirclles", and the phenomenon of their appearance as "snircling". In order to quantify their appearance, a method had to be found to isolate them as objects in the image field, and then enumerate them, measure their area, or determine some parameter about them which correlated in a dose-dependent fashion with the concentration of insulin to which the cells had been exposed. In order to separate the snirclles, a binarization procedure was applied in which one copy of the image smoothed with a relatively severe gaussian kernel ($\sigma = 2.5$) was subtracted from another copy to which only a relatively light gaussian smooth had been applied ($\sigma = 0.5$). The resultant image was rescaled to its min/max range, and an automatic threshold was applied to divide the image into two levels. The thresholded image contains a background of one value all found object with another value. The found objects were first filtered through a filter to remove objects far too large and far too small to be snirclles. The remaining objects, which represent snirclles and other artifacts from the image with approximately the same size and intensity characteristics as snirclles, are passed into a classification procedure which has been previously trained with many images of snirclles to recognize snirclles and exclude the

other artifacts. The result of this procedure is a binary image that shows only the found snircles to the degree to which the classification procedure can accurately identify them. The total area of the snircles is then summed and this value is the quantitative measure of the degree of snircling for that image.

- 5 Another approach to the problem of quantifying GLUT 4 translocation has been performed and validated using the same type of experimental protocol but a different image processing approach. In this case the objects of interest in the cells are not the appearance of snircles at the plasma membrane but the disappearance of GLUT4-GFP fluorescence from its intracellular site. With this method the bright area, consisting of GLUT4-GFP, 10 centrally located in each cell is identified by a thresholding procedure. This demarcates a certain area for the centrally located GLUT4-GFP. In the next step the total fluorescence intensity in this area is quantified on each image in the image series, i.e. over time. The response for each cell is defined as the difference in fluorescence intensity in the centrally located GLUT4-GFP area before and a fixed point in time after application of the 15 influence. The dose-response relationship for insulin using the above described quantitation procedure is shown in Figure 13. It can be seen that the ED50 value for insulin to reduce central GLUT4-GFP fluorescence is 0.3 nM.

In the present specification and claims, the term "an influence" covers any influence to which the cellular response comprises a redistribution. Thus, e.g., heating, cooling, high 20 pressure, low pressure, humidifying, or drying are influences on the cellular response on which the resulting redistribution can be quantified, but as mentioned above, perhaps the most important influences are the influences of contacting or incubating the cells with substances which are known or suspected to exert an influence on the cellular response involving a redistribution contribution. In another embodiment of the invention the 25 influence could be substances from a compound drug library.

In the present context, the term "green fluorescent protein" is intended to indicate a protein which, when expressed by a cell, emits fluorescence upon exposure to light of the correct excitation wavelength (cf. [(Chalfie, M. *et al.* (1994) Science 263, 802-805])). In the following, GFP in which one or more amino acids have been substituted, inserted or 30 deleted is most often termed "modified GFP". "GFP" as used herein includes wild-type

GFP derived from the jelly fish *Aequorea victoria* and modifications of GFP, such as the blue fluorescent variant of GFP disclosed by Heim *et al.* (1994). Proc.Natl.Acad.Sci. 91:26, pp 12501-12504, and other modifications that change the spectral properties of the GFP fluorescence, or modifications that exhibit increased fluorescence when expressed in cells
5 at a temperature above about 30°C described in PCT/DK96/00051, published as WO 97/11094 on 27 March 1997 and hereby incorporated by reference, and which comprises a fluorescent protein derived from *Aequorea* Green Fluorescent Protein (GFP) or any functional analogue thereof, wherein the amino acid in position 1 upstream from the chromophore has been mutated to provide an increase of fluorescence intensity when the fluorescent protein of
10 the invention is expressed in cells. Preferred GFP variants are F64L-GFP, F64L-Y66H-GFP and F64L-S65T-GFP. An especially preferred variant of GFP for use in all the aspects of this invention is EGFP (DNA encoding EGFP which is a F64L-S65T variant with codons optimized for expression in mammalian cells is available from Clontech, Palo Alto, plasmids containing the EGFP DNA sequence, cf. GenBank Acc. Nos. U55762, U55763).

15 The term "intracellular signalling pathway" and "signal transduction pathway" are intended to indicate the co-ordinated intracellular processes whereby a living cell transduce an external or internal signal into cellular responses. Said signal transduction will involve an enzymatic reaction said enzymes include but are not limited to protein kinases, GTPases, ATPases, protein phosphatases, phospholipases and cyclic nucleotide
20 phosphodiesterases. The cellular responses include but are not limited to gene transcription, secretion, proliferation, mechanical activity, metabolic activity, cell death.

The term "second messenger" is used to indicate a low molecular weight component involved in the early events of intracellular signal transduction pathways.

The term "luminophore" is used to indicate a chemical substance that has the property of
25 emitting light either inherently or upon stimulation with chemical or physical means. This includes but is not limited to fluorescence, bioluminescence, phosphorescence, and chemiluminescence.

The term "mechanically intact living cell" is used to indicate a cell which is considered living according to standard criteria for that particular type of cell such as maintenance of
30 normal membrane potential, energy metabolism, proliferative capability, and has not

experienced any physically invasive treatment designed to introduce external substances into the cell such as microinjection.

In the present context, the term "permeabilised living cell" is used to indicate cells where a pore forming agent such as Streptolysin O or *Staphylococcus Aureus* α -toxin has been applied and thereby incorporated into the plasma membrane in the cells. This creates proteinaceous pores with a defined pore size in the plasma membranes of the exposed cells. Pores could also be made by electroporation, i.e. exposing the cells to high voltage discharges, a procedure that creates small holes in the plasma membrane by coagulating integral membrane proteins. Treatment with a mild detergent such as saponin may accomplish the same thing. Common to all these treatments are that pores are formed only in the plasma membrane without affecting the integrity of cytoplasmic structural elements and organelles. The term living in this context means that the permeabilised cells bathed in a solution mimicking the intracellular milieu still have functional organelles, such as actively respiring mitochondria and endoplasmic reticulum that can take up and release calcium ions, and functional structural elements. The benefit of this method is that substances that normally can not traverse the plasma membrane, but most likely exert their influence intracellularly, can be introduced and their influence studied without cumbersome microinjection of the substances into single cells. Using this method the response to an influence can be recorded from many cells simultaneously.

In the present context, the term "permeabilisation" is intended to indicate the selective disruption of the plasma membrane barrier so that soluble substances freely mobile in the cytosol are lost from the cells. The permeabilisation can be achieved as described above under "permeabilised living cells" or by using other chemical detergents such as Triton X-100 or digitonin in carefully titrated amounts.

The term "physiologically relevant", when applied to an experimentally determined redistribution of an intracellular component, as measured by a change in the luminescence properties or distribution, is used to indicate that said redistribution can be explained in terms of the underlying biological phenomenon which gives rise to the redistribution.

The terms "image processing" and "image analysis" are used to describe a large family of digital data analysis techniques or combination of such techniques which reduce ordered

arrays of numbers (images) to quantitative information describing those ordered arrays of numbers. When said ordered arrays of numbers represent measured values from a physical process, the quantitative information derived is therefore a measure of the physical process.

- 5 The term "fluorescent probe" is used to indicate a fluorescent fusion polypeptide comprising a GFP or any functional part thereof which is N- or C-terminally fused to a biologically active polypeptide as defined herein, optionally via a peptide linker consisting of one or more amino acid residues, where the size of the linker peptide in itself is not critical as long as the desired functionality of the fluorescent probe is maintained. A
- 10 fluorescent probe according to the invention is expressed in a cell and basically mimics the physiological behaviour of the biologically active polypeptide moiety of the fusion polypeptide.

The term "mammalian cell" is intended to indicate any living cell of mammalian origin. The cell may be an established cell line, many of which are available from The American

15 Type Culture Collection (ATCC, Virginia, USA) or a primary cell with a limited life span derived from a mammalian tissue, including tissues derived from a transgenic animal, or a newly established immortal cell line derived from a mammalian tissue including transgenic tissues, or a hybrid cell or cell line derived by fusing different cell types of mammalian origin e.g. hybridoma cell lines. The cells may optionally express one or more

20 non-native gene products, e.g. receptors, enzymes, enzyme substrates, prior to or in addition to the fluorescent probe. Preferred cell lines include but are not limited to those of fibroblast origin, e.g. BHK, CHO, BALB, or of endothelial origin, e.g. HUVEC, BAE (bovine artery endothelial), CPAE (cow pulmonary artery endothelial), HLMVEC (human lung microvascular endothelial cells) or of pancreatic origin, e.g. RIN, INS-1, MIN6,

25 bTC3, aTC6, bTC6, HIT, or of hematopoietic origin, e.g. primary isolated human monocytes, macrophages, neutrophils, basophils, eosinophils and lymphocyte populations, AML-193, HL-60, RBL-1, adipocyte origin, e.g. 3T3-L1, neuronal/neuroendocrine origin, e.g. AtT20, PC12, GH3, muscle origin, e.g. SKMC, A10, C2C12, renal origin, e.g. HEK 293, LLC-PK1.

- 30 The term "hybrid polypeptide" is intended to indicate a polypeptide which is a fusion of at

least a portion of each of two proteins, in this case at least a portion of the green fluorescent protein, and at least a portion of a catalytic and/or regulatory domain of a protein kinase. Furthermore a hybrid polypeptide is intended to indicate a fusion polypeptide comprising a GFP or at least a portion of the green fluorescent protein that
5 contains a functional fluorophore, and at least a portion of a biologically active polypeptide as defined herein provided that said fusion is not the PKC α -GFP, PKC γ -GFP, and PKC ϵ -GFP disclosed by Schmidt *et al.* and Sakai *et al.*, respectively. Thus, GFP may be N- or C-terminally tagged to a biologically active polypeptide, optionally via a linker portion or linker peptide consisting of a sequence of one or more amino acids. The hybrid
10 polypeptide or fusion polypeptide may act as a fluorescent probe in intact living cells carrying a DNA sequence encoding the hybrid polypeptide under conditions permitting expression of said hybrid polypeptide.

The term "kinase" is intended to indicate an enzyme that is capable of phosphorylating a cellular component.

15 The term "protein kinase" is intended to indicate an enzyme that is capable of phosphorylating serine and/or threonine and/or tyrosine in peptides and/or proteins.

The term "phosphatase" is intended to indicate an enzyme that is capable of dephosphorylating phosphoserine and/or phosphothreonine and/or phosphotyrosine in peptides and/or proteins.

20 The term "cyclic nucleotide phosphodiesterase" is intended to indicate an enzyme that is capable of inactivating the second messengers cAMP and cGMP by hydrolysis of their 3'-ester bond.

In the present context, the term "biologically active polypeptide" is intended to indicate a polypeptide affecting intracellular processes upon activation, such as an enzyme which is
25 active in intracellular processes or a portion thereof comprising a desired amino acid sequence which has a biological function or exerts a biological effect in a cellular system. In the polypeptide one or several amino acids may have been deleted, inserted or replaced to alter its biological function, e.g. by rendering a catalytic site inactive. Preferably, the biologically active polypeptide is selected from the group consisting of proteins taking part

in an intracellular signalling pathway, such as enzymes involved in the intracellular phosphorylation and dephosphorylation processes including kinases, protein kinases and phosphorylases as defined herein, but also proteins making up the cytoskeleton play important roles in intracellular signal transduction and are therefore included in the meaning of "biologically active polypeptide" herein. More preferably, the biologically active polypeptide is a protein which according to its state as activated or non-activated changes localisation within the cell, preferably as an intermediary component in a signal transduction pathway. Included in this preferred group of biologically active polypeptides are cAMP dependent protein kinase A.

10 The term "a substance having biological activity" is intended to indicate any sample that has a biological function or exerts a biological effect in a cellular system. The sample may be a sample of a biological material such as a sample of a body fluid including blood, plasma, saliva, milk, urine, or a microbial or plant extract, an environmental sample containing pollutants including heavy metals or toxins, or it may be a sample containing a
15 compound or mixture of compounds prepared by organic synthesis or genetic techniques.

The phrase "any change in fluorescence" means any change in absorption properties, such as wavelength and intensity, or any change in spectral properties of the emitted light, such as a change of wavelength, fluorescence lifetime, intensity or polarisation, or any change in the intracellular localisation of the fluorophore. It may thus be localised to a specific
20 cellular component (e.g. organelle, membrane, cytoskeleton, molecular structure) or it may be evenly distributed throughout the cell or parts of the cell.

The term "organism" as used herein indicates any unicellular or multicellular organism preferably originating from the animal kingdom including protozoans, but also organisms that are members of the plant kingdoms, such as algae, fungi, bryophytes, and vascular
25 plants are included in this definition.

The term "nucleic acid" is intended to indicate any type of poly- or oligonucleic acid sequence, such as a DNA sequence, a cDNA sequence, or an RNA sequence.

The term "biologically equivalent" as it relates to proteins is intended to mean that a first protein is equivalent to a second protein if the cellular functions of the two proteins may

substitute for each other, e.g. if the two proteins are closely related isoforms encoded by different genes, if they are splicing variants, or allelic variants derived from the same gene, if they perform identical cellular functions in different cell types, or in different species. The term "biologically equivalent" as it relates to DNA is intended to mean that a first
5 DNA sequence encoding a polypeptide is equivalent to a second DNA sequence encoding a polypeptide if the functional proteins encoded by the two genes are biologically equivalent.

The phrase "back-tracking of a signal transduction pathway" is intended to indicate a process for defining more precisely at what level a signal transduction pathway is affected,
10 either by the influence of chemical compounds or a disease state in an organism. Consider a specific signal transduction pathway represented by the bioactive polypeptides A - B - C - D, with signal transduction from A towards D. When investigating all components of this signal transduction pathway compounds or disease states that influence the activity or redistribution of only D can be considered to act on C or downstream of C whereas
15 compounds or disease states that influence the activity or redistribution of C and D, but not of A and B can be considered to act downstream of B.

The term "fixed cells" is used to mean cells treated with a cytological fixative such as glutaraldehyde or formaldehyde, treatments that serve to chemically cross-link and stabilise soluble and insoluble proteins within the structure of the cell. Once in this state,
20 such proteins cannot be lost from the structure of the now-dead cell.

In the present context a "screening assay" is intended to mean any measurement protocol, including materials, cells, instruments, chemicals, reagents, detection units, calibration and quantification procedures used to measure a response from mechanically intact or permeabilised living cells relevant to influences on an intracellular pathway.

25 The term "dose-response relationship" and "screening programme" is in the present context intended to mean a clear correlation between the quantified response of cells in a screening assay to application of an influence, such as a compound, and the concentration of the applied influence. The response to the influence may be both an up-regulation and a down-regulation of the quantified parameter used in the screening assay.

In the present context, the term "physiology" is intended to mean the normal function of biological and biochemical processes inside cells, between cells and in the whole organism or animal.

5 BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. CHO cells expressing the PKAc-F64L-S65T-GFP hybrid protein have been treated in HAM's F12 medium with 50 μ M forskolin at 37°C. The images of the GFP fluorescence in these cells have been taken at different time intervals after treatment, which were: a) 40 seconds b) 60 seconds c) 70 seconds d) 80 seconds. The fluorescence
10 changes from a punctate to a more even distribution within the (non-nuclear) cytoplasm.

Figure 2. Time-lapse analysis of forskolin induced PKAc-F64L-S65T-GFP redistribution. CHO cells, expressing the PKAc-F64L-S65T-GFP fusion protein were analysed by time-lapse fluorescence microscopy. Fluorescence micrographs were acquired at regular
15 intervals from 2 min before to 8 min after the addition of agonist. The cells were challenged with 1 μ M forskolin immediately after the upper left image was acquired (t=0). Frames were collected at the following times: i) 0, ii) 1, iii) 2, iv) 3, v) 4 and vi) 5 minutes. Scale bar 10 μ m.

20 Figure 3. Time-lapse analyses of PKAc-F64L-S65T-GFP redistribution in response to various agonists. The effects of 1 μ M forskolin (A), 50 μ M forskolin (B), 1mM dbcAMP (C) and 100 μ M IBMX (D) (additions indicated by open arrows) on the localisation of the PKAc-F64L-S65T-GFP fusion protein were analysed by time-lapse fluorescence microscopy of CHO/PKAc-F64L-S65T-GFP cells. The effect of addition of 10 μ M
25 forskolin (open arrow), followed shortly by repeated washing with buffer (solid arrow), on the localisation of the PKAc-F64L-S65T-GFP fusion protein was analysed in the same cells (E). In a parallel experiment, the effect of adding 10 μ M forskolin and 100 μ M IBMX (open arrow) followed by repeated washing with buffer containing 100 μ M IBMX

(solid arrow) was analysed (F). Removing forskolin caused PKAc-F64L-S65T-GFP fusion protein to return to the cytoplasmic aggregates while this is prevented by the continued presence of IBMX (F). The effect of 100 nM glucagon (Fig 3G, open arrow) on the localisation of the PKAc-F64L-S65T-GFP fusion protein is also shown for BHK/GR, PKAc-F64L-S65T-GFP cells. The effect of 10 μ M norepinephrine (H), solid arrow, on the localisation of the PKAc-F64L-S65T-GFP fusion protein was analysed similarly, in transiently transfected CHO, PKAc-F64L-S65T-GFP cells, pretreated with 10 μ M forskolin, open arrow, to increase [cAMP]. N.B. in Fig 3H the x-axis counts the image numbers, with 12 seconds between images. The raw data of each experiment consisted of 60 fluorescence micrographs acquired at regular intervals including several images acquired before the addition of buffer or agonist. The charts (A-G) each show a quantification of the response seen through all the 60 images, performed as described in analysis method 2. The change in total area of the highly fluorescent aggregates, relative to the initial area of fluorescent aggregates is plotted as the ordinate in all graphs in Figure 3, versus time for each experiment. Scale bar 10 μ m.

Figure 4. Dose-response curve (two experiments) for forskolin-induced redistribution of the PKAc-F64L-S65T-GFP fusion.

Figure 5. Time from initiation of a response to half maximal ($t_{1/2\max}$) and maximal (t_{\max}) PKAc-F64L-S65T-GFP redistribution. The data was extracted from curves such as that shown in "Figure 2." All $t_{1/2\max}$ and t_{\max} values are given as mean \pm SD and are based on a total of 26-30 cells from 2-3 independent experiments for each forskolin concentration. Since the observed redistribution is sustained over time, the t_{\max} values were taken as the earliest time point at which complete redistribution is reached. Note that the values do not relate to the degree of redistribution.

Figure 6. Parallel dose-response analyses of forskolin induced cAMP elevation and PKAc-

F64L-S65T-GFP redistribution. The effects of buffer or 5 increasing concentrations of forskolin on the localisation of the PKAc-F64L-S65T-GFP fusion protein in CHO/PKAc-F64L-S65T-GFP cells, grown in a 96 well plate, were analysed as described above. Computing the ratio of the SD's of fluorescence micrographs taken of the same field of
5 cells, prior to and 30 min after the addition of forskolin, gave a reproducible measure of PKAc-F64L-S65T-GFP redistribution. The graph shows the individual 48 measurements and a trace of their mean \pm s.e.m at each forskolin concentration. For comparison, the effects of buffer or 8 increasing concentrations of forskolin on $[cAMP]_i$ was analysed by a scintillation proximity assay of cells grown under the same conditions. The graph shows a
10 trace of the mean \pm s.e.m of 4 experiments expressed in arbitrary units.

Figure 7. BHK cells stably transfected with the human muscarinic (hM1) receptor and the PKC α -F64L-S65T-GFP fusion. Carbachol (100 μ M added at 1.0 second) induced a transient redistribution of PKC α -F64L-S65T-GFP from the cytoplasm to the plasma
15 membrane. Images were taken at the following times: a) 1 second before carbachol addition, b) 8.8 seconds after addition and c) 52.8 seconds after addition.

Figure 8. BHK cells stably transfected with the hM1 receptor and PKC α -F64L-S65T-GFP fusion were treated with carbachol (1 μ M, 10 μ M, 100 μ M). In single cells intracellular
20 $[Ca^{2+}]$ was monitored simultaneously with the redistribution of PKC α -F64L-S65T-GFP. Dashed line indicates the addition times of carbachol. The top panel shows changes in the intracellular Ca^{2+} concentration of individual cells with time for each treatment. The middle panel shows changes in the average cytoplasmic GFP fluorescence for individual cells against time for each treatment. The bottom panel shows changes in the fluorescence
25 of the periphery of single cells, within regions that specifically include the circumferential edge of a cell as seen in normal projection, the best regions for monitoring changes in the fluorescence intensity of the plasma membrane.

Figure 9.

- a) The hERK1-F64L-S65T-GFP fusion expressed in HEK293 cells treated with 100 μ M of the MEK1 inhibitor PD98059 in HAM F-12 (without serum) for 30 minutes at 37 °C. The nuclei empty of fluorescence during this treatment.
- 5 b) The same cells as in (a) following treatment with 10 % foetal calf serum for 15 minutes at 37 °C.
- c) Time profiles for the redistribution of GFP fluorescence in HEK293 cells following treatment with various concentrations of EGF in Hepes buffer (HAM F-12 replaced with Hepes buffer directly before the experiment). Redistribution of fluorescence is expressed as the change in the ratio value between areas in nucleus and cytoplasm of single cells. Each time profile is the mean for the changes seen in six single cells.
- 10 d) Bar chart for the end-point measurements, 600 seconds after start of EGF treatments, of fluorescence change (nucleus:cytoplasm) following various concentrations of EGF.

15 Figure 10.

- a) The SMAD2-EGFP fusion expressed in HEK293 cells starved of serum overnight in HAM F-12. HAM F-12 was then replaced with Hepes buffer pH 7.2 immediately before the experiment. Scale bar is 10 μ m.
- b) HEK 293 cells expressing the SMAD2-EGFP fusion were treated with various concentration of TGF-beta as indicated, and the redistribution of fluorescence monitored against time. The time profile plots represent increases in fluorescence within the nucleus, normalised to starting values in each cell measured. Each trace is the time profile for a single cell nucleus.
- 20 c) A bar chart representing the end-point change in fluorescence within nuclei (after 850 seconds of treatment) for different concentrations of TGF-beta. Each bar is the value for a single nucleus in each treatment.
- 25

Figure 11. The VASP-F64L-S65T-GFP fusion in CHO cells stably transfected with the human insulin receptor. The cells were starved for two hours in HAM F-12 without serum, then treated with 10% foetal calf serum. The image shows the resulting redistribution of
5 fluorescence after 15 minutes of treatment. GFP fluorescence becomes localised in structures identified as focal adhesions along the length of actin stress fibres.

Figure 12. Time lapse recording GLUT4-GFP redistribution in CHO-HIR cells. Time indicates minutes after the addition of 100 nM insulin.

10

Figure 13. Dose-response relationships for the influence of insulin on the disappearance of total fluorescence from the centrally located area of GLUT4-GFP. Data points indicate mean \pm SE.

15 Figure 14. Dose-response relationship for the translocation of PKC α -GFP in BHKhM1 cells stimulated with the muscarinic agonist carbamylcholine using a FLIPRTM to do the actual experiments.

Figure 15. Dose-response relationship for the translocation of PKAc-GFP in CHO/PKAc-F64L-S65T-GFP cells stimulated with forskolin using a FLIPRTM to do the actual
20 experiments.

Figure 16. Dose-response relationship for the disappearance of fluorescence from permeabilised CHO/PKAc-F64L-S65T-GFP when previously exposed to different doses of
25 forskolin.

EXAMPLES

EXAMPLE 1

Construction, testing and implementation of an assay for cAMP based on PKA 5 activation in real time within living cells.

Useful for monitoring the activity of signalling pathways that lead to altered concentrations of cAMP, e.g. activation of G-protein coupled receptors which couple to G-proteins of the G_s or G_i class.

The catalytic subunit of the murine cAMP dependent protein kinase (PKAc) was fused C-terminally to a F64L-S65T derivative of GFP. The resulting fusion (PKAc-F64L-S65T-
10 GFP) was used for monitoring *in vivo* the translocation and thereby the activation of PKA.

To construct the PKAc-F64L-S65T-GFP fusion, convenient restriction endonuclease sites were introduced into the cDNAs encoding murine PKAc (Gen Bank Accession number: M12303) and F64L-S65T-GFP (sequence disclosed in WO 97/11094) by polymerase chain
15 reaction (PCR). The PCR reactions were performed according to standard protocols with the following primers:

5'PKAc:

TTggACACAAgCTTTggACACCCTCAggATATgggCAACgCCgCCgCCgCCAAg (SEQ
ID NO:3),

20 3'PKAc:

gTCATCTTCTCgAgTCTTTCAGgCgCgCCCAAACCTCAgTAAACTCCTTgCCACAC
(SEQ ID NO:4) ,

5'GFP: TTggACACAAgCTTTggACACggCgCgCCATgAgTAAAggAgAAgAACTTTTC
(SEQ ID NO:1),

25 3'GFP: gTCATCTTCTCgAgTCTTACTCCTgAggTTTgTATAgTTCATCCATgCCATgT
(SEQ ID NO:2).

The PKAc amplification product was then digested with HindIII+AscI and the F64L-S65T-GFP product with AscI+XhoI. The two digested PCR products were subsequently ligated with a HindIII+XhoI digested plasmid (pZeoSV® mammalian expression vector, Invitrogen, San Diego, CA, USA). The resulting fusion construct (SEQ ID NO:68 & 69) was under control of the SV40 promoter.

Transfection and cell culture conditions:

Chinese hamster ovary cells (CHO), were transfected with the plasmid containing the PKAc-F64L-S65T-GFP fusion using the calcium phosphate precipitate method in HEPES-buffered saline (Sambrook *et al.*, 1989). Stable transfectants were selected using 1000 µg Zeocin/ml (Invitrogen) in the growth medium (DMEM with 1000 mg glucose/l, 10 % fetal bovine serum (FBS), 100 µg penicillin-streptomycin mixture ml⁻¹, 2 mM L-glutamine purchased from Life Technologies Inc., Gaithersburg, MD, USA). Untransfected CHO cells were used as the control. To assess the effect of glucagon on fusion protein translocation, the PKAc-F64L-S65T-GFP fusion was stably expressed in baby hamster kidney cells overexpressing the human glucagon receptor (BHK/GR cells). Untransfected BHK/GR cells were used as the control. Expression of GR was maintained with 500 µg G418/ml (*Neo* marker) and PKAc-F64L-S65T-GFP was maintained with 500 µg Zeocin/ml (*Sh ble* marker). CHO cells were also simultaneously co-transfected with vectors containing the PKAc-F64L-S65T-GFP fusion and the human α2a adrenoceptor (hARα2a).

For fluorescence microscopy, cells were allowed to adhere to Lab-Tek chambered coverglasses (Nalge Nunc Int., Naperville, IL, USA) for at least 24 hours and cultured to about 80% confluence. Prior to experiments, the cells were cultured over night without selection pressure in HAM F-12 medium with glutamax (Life Technologies), 100 µg penicillin-streptomycin mixture ml⁻¹ and 0.3 % FBS. This medium has low autofluorescence enabling fluorescence microscopy of cells straight from the incubator.

Monitoring activity of PKA activity in real time:

Image acquisition of live cells were gathered using a Zeiss Axiovert 135M fluorescence microscope fitted with a Fluar 40X, NA: 1.3 oil immersion objective and coupled to a Photometrics CH250 charged coupled device (CCD) camera. The cells were illuminated with a 100 W HBO arc lamp. In the light path was a 470 ± 20 nm excitation filter, a 510 nm dichroic mirror and a 515 ± 15 nm emission filter for minimal image background. The cells were maintained at 37°C with a custom built stage heater.

Images were processed and analysed in the following manner:

Method 1: Stepwise procedure for quantitation of translocation of PKA:

1. The image was corrected for dark current by performing a pixel-by-pixel subtraction of a dark image (an image taken under the same conditions as the actual image, except the camera shutter is not allowed to open).
2. The image was corrected for non-uniformity of the illumination by performing a pixel-by-pixel ratio with a flat field correction image (an image taken under the same conditions as the actual image of a uniformly fluorescent specimen).
3. The image histogram, i.e., the frequency of occurrence of each intensity value in the image, was calculated.
4. A smoothed, second derivative of the histogram was calculated and the second zero is determined. This zero corresponds to the inflection point of the histogram on the high side of the main peak representing the bulk of the image pixel values.
5. The value determined in step 4 was subtracted from the image. All negative values were discarded.
6. The variance (square of the standard deviation) of the remaining pixel values was determined. This value represents the "response" for that image.
7. Scintillation proximity assay (SPA) for independent quantitation of cAMP.

Method 2: Alternative method for quantitation of PKA redistribution:

1. The fluorescent aggregates are segmented from each image using an automatically found threshold based on the maximisation of the information measure between the object and background. The *a priori* entropy of the image histogram is used as the information measure.
2. The area of each image occupied by the aggregates is calculated by counting pixels in the segmented areas.
3. The value obtained in step 2 for each image in a series, or treatment pair, is normalised to the value found for the first (unstimulated) image collected. A value of zero (0) indicates no redistribution of fluorescence from the starting condition. A value of one (1) by this method equals full redistribution.

Cells were cultured in HAM F-12 medium as described above, but in 96-well plates. The medium was exchanged with Ca^{2+} -HEPES buffer including 100 μM IBMX and the cells were stimulated with different concentrations of forskolin for 10 min. Reactions were stopped with addition of NaOH to 0.14 M and the amount of cAMP produced was measured with the cAMP-SPA kit, RPA538 (Amersham) as described by the manufacturer.

Manipulating intracellular levels of cAMP to test the PKAc-F64L-S65T-GFP fusion.

The following compounds were used to vary cAMP levels: Forskolin, an activator of adenylate cyclase; dbcAMP, a membrane permeable cAMP analog which is not degraded by phosphodiesterase; IBMX, an inhibitor of phosphodiesterase.

CHO cells stably expressing the PKAc-F64L-S65T-GFP, showed a dramatic translocation of the fusion protein from a punctate distribution to an even distribution throughout the cytoplasm following stimulation with 1 μM forskolin ($n=3$), 10 μM forskolin ($n=4$) and 50 μM forskolin ($n=4$) (Fig 1), or dbcAMP at 1mM ($n=6$).

Fig. 2 shows the progression of response in time following treatment with 1 μM forskolin.

Fig. 3 gives a comparison of the average temporal profiles of fusion protein redistribution and a measure of the extent of each response to the three forskolin concentrations (Fig. 3A, E, B), and to 1 mM dbcAMP (fig 3C) which caused a similar but slower response, and to addition of 100 μ M IBMX (n=4, Fig. 3D) which also caused a slow response, even in the
5 absence of adenylate cyclase stimulation. Addition of buffer (n=2) had no effect (data not shown).

As a control for the behaviour of the fusion protein, F64L-S65T-GFP alone was expressed in CHO cells and these were also given 50 μ M forskolin (n=5); the uniform diffuse
10 distribution characteristic of GFP in these cells was unaffected by such treatment (data not shown).

The forskolin-induced translocation of PKAc-F64L-S65T-GFP showed a dose-response relationship (Fig 4 and 6), see quantitative procedures above.

Reversibility of PKAc-F64L-S65T-GFP translocation.

15 The release of the PKAc probe from its cytoplasmic anchoring hotspots was reversible. Washing the cells repeatedly (5-8 times) with buffer after 10 μ M forskolin treatment completely restored the punctate pattern within 2-5 min (n=2, Fig. 3E). In fact the fusion protein returned to a pattern of fluorescent cytoplasmic aggregates virtually indistinguishable from that observed before forskolin stimulation.

20 To test whether the return of fusion protein to the cytoplasmic aggregates reflected a decreased [cAMP], cells were treated with a combination of 10 μ M forskolin and 100 μ M IBMX (n=2) then washed repeatedly (5-8 times) with buffer containing 100 μ M IBMX (Fig. 3F). In these experiments, the fusion protein did not return to its prestimulatory localisation after removal of forskolin.

25

Testing the PKA-F64L-S65T-GFP probe with physiologically relevant agents.

- To test the probe's response to receptor activation of adenylate cyclase, BHK cells stably transfected with the glucagon receptor and the PKA-F64L-S65T-GFP probe were exposed to glucagon stimulation. The glucagon receptor is coupled to a G_s protein which activates adenylate cyclase, thereby increasing the cAMP level. In these cells, addition of 100 nM glucagon (n=2) caused the release of the PKA-F64L-S65T-GFP probe from the cytoplasmic aggregates and a resulting translocation of the fusion protein to a more even cytoplasmic distribution within 2-3 min (Fig. 3G). Similar but less pronounced effects were seen at lower glucagon concentrations (n=2, data not shown). Addition of buffer (n=2) had no effect over time (data not shown).
- 10 Transiently transfected CHO cells expressing hAR α 2a and the PKA-F64L-S65T-GFP probe were treated with 10 μ M forskolin for 7.5 minutes, then, in the continued presence of forskolin, exposed to 10 μ M norepinephrine to stimulate the exogenous adrenoreceptors, which couple to a G_i protein, which inhibit adenylate cyclase. This treatment led to reappearance of fluorescence in the cytoplasmic aggregates indicative of a decrease in [cAMP]_i (Fig. 3H).
- 15

Fusion protein translocation correlated with [cAMP]_i

- As described above, the time it took for a response to come to completion was dependent on the forskolin dose (Fig. 5) In addition the degree of responses was also dose-dependent.
- 20 To test the PKA-F64L-S65T-GFP fusion protein translocation in a semi high through-put system, CHO cells stably transfected with the PKA-F64L-S65T-GFP fusion was stimulated with buffer and 5 increasing doses of forskolin (n=8). Using the image analysis algorithm described above (Method 1), a dose-response relationship was observed in the range from 0.01-50 μ M forskolin (Fig. 6). A half-maximal stimulation was observed at about 2 μ M forskolin. In parallel, cells were stimulated with buffer and 8 increasing concentrations of forskolin (n=4) in the range 0.01-50 μ M. The amount of cAMP produced was measured in an SPA assay. A steep increase was observed between 1 and 5 μ M forskolin coincident with the steepest part of the curve for fusion protein translocation (also Fig. 6).
- 25

EXAMPLE 2

Quantitation of redistribution in real-time within living cells.

Probe for detection of PKC activity in real time within living cells:

5 Construction of PKC-GFP fusion:

The probe was constructed by ligating two restriction enzyme treated polymerase chain reaction (PCR) amplification products of the cDNA for murine PKC α (GenBank Accession number: M25811) and F64L-S65T-GFP (sequence disclosed in WO 97/11094) respectively. Taq® polymerase and the following oligonucleotide primers were used for
10 PCR;

5'mPKC α :

TTggACACAAgCTTTggACACCCTCAggATATggCTgACgTTTACCCggCCAACg
(SEQ ID NO:5),

3'mPKC α :

15 gTCATCTTCTCgAgTCTTTCAGgCgCgCCCTACTgCACTTTgCAAgATTgggTgC (SEQ ID NO:6),

5'F64L-S65T-GFP:

TTggACACAAgCTTTggACACggCgCgCCATgAgTAAAggAgAAgAACTTTTC (SEQ ID NO:1),

20 3'F64L-S65T-GFP:

gTCATCTTCTCgAgTCTTACTCCTgAggTTTgTATAgTTCATCCATgCCATgT (SEQ ID NO:2).

The hybrid DNA strand was inserted into the pZeoSV® mammalian expression vector as a HindIII-XhoI cassette as described in example 1.

25 BHK cells expressing the human M1 receptor under the control of the inducible metallothionine promoter and maintained with the dihydrofolate reductase marker were

transfected with the PKC α -F64L-S65T-GFP probe using the calcium phosphate precipitate method in HEPES buffered saline (HBS [pH 7.10]). Stable transfectants were selected using 1000 μ g Zeocin®/ml in the growth medium (DMEM with 1000 mg glucose/l, 10 % foetal bovine serum (FBS), 100 μ g penicillin-streptomycin mixture ml⁻¹, 2 mM l-glutamine). The hM1 receptor and PKC α -F64L-S65T-GFP fusion protein were maintained with 500 nM methotrexate and 500 μ g Zeocin®/ml respectively. 24 hours prior to any experiment, the cells were transferred to HAM F-12 medium with glutamax, 100 μ g penicillin-streptomycin mixture ml⁻¹ and 0.3 % FBS. This medium relieves selection pressure, gives a low induction of signal transduction pathways and has a low autofluorescence at the relevant wavelength enabling fluorescence microscopy of cells straight from the incubator.

Method 1: Monitoring the PKC α activity in real time:

Digital images of live cells were gathered using a Zeiss Axiovert 135M fluorescence microscope fitted with a 40X, NA: 1.3 oil immersion objective and coupled to a Photometrics CH250 charged coupled device (CCD) camera. The cells were illuminated with a 100 W arc lamp. In the light path was a 470 \pm 20 nm excitation filter, a 510 nm dichroic mirror and a 515 \pm 15 nm emission filter for minimal image background. The cells were kept and monitored to be at 37°C with a custom built stage heater.

Images were analyzed using the IPLab software package for Macintosh.

Upon stimulation of the M1-BHK cells, stably expressing the PKC α -F64L-S65T-GFP fusion, with carbachol we observed a dose-dependent transient translocation from the cytoplasm to the plasma membrane (Fig. 7a,b,c). Simultaneous measurement of the cytosolic free calcium concentration shows that the carbachol-induced calcium mobilisation precedes the translocation (Fig. 8).

Stepwise procedure for quantitation of translocation of PKC α :

1. The image was corrected for dark current by performing a pixel-by-pixel subtraction of a dark image (an image taken under the same conditions as the actual image, except the camera shutter is not allowed to open).
2. The image was corrected for non-uniformity of the illumination by performing a pixel-by-pixel ratio with a flat field correction image (an image taken under the same conditions as the actual image of a uniformly fluorescent specimen).
3. A copy of the image was made in which the edges are identified. The edges in the image are found by a standard edge-detection procedure – convolving the image with a kernel which removes any large-scale unchanging components (i.e., background) and accentuates any small-scale changes (i.e., sharp edges). This image was then converted to a binary image by thresholding. Objects in the binary image which are too small to represent the edges of cells were discarded. A dilation of the binary image was performed to close any gaps in the image edges. Any edge objects in the image which were in contact with the borders of the image are discarded. This binary image represents the edge mask.
4. Another copy of image was made via the procedure in step 3. This copy was further processed to detect objects which enclose “holes” and setting all pixels inside the holes to the binary value of the edge, i.e., one. This image represents the whole cell mask.
5. The original image was masked with the edge mask from step 3 and the sum total of all pixel values is determined.
6. The original image was masked with the whole cell mask from step 4 and the sum total of all pixel values was determined.
7. The value from step 5 was divided by the value from step 6 to give the final result, the fraction of fluorescence intensity in the cells which was localized in the edges.

EXAMPLE 3

Probes for detection of mitogen activated protein kinase Erk1 redistribution.

Useful for monitoring signalling pathways involving MAPK, e.g. to identify compounds which modulate the activity of the pathway in living cells.

Erk1, a serine/threonine protein kinase, is a component of a signalling pathway that is
5 activated by e.g. many growth factors.

Probes for detection of ERK-1 activity in real time within living cells:

The extracellular signal regulated kinase (ERK-1, a mitogen activated protein kinase, MAPK) is fused N- or C-terminally to a derivative of GFP. The resulting fusions
10 expressed in different mammalian cells are used for monitoring *in vivo* the nuclear translocation, and thereby the activation, of ERK1 in response to stimuli that activate the MAPK pathway.

a) Construction of murine ERK1 - F64L-S65T-GFP fusion:

Convenient restriction endonuclease sites are introduced into the cDNAs encoding murine ERK1 (GenBank Accession number: Z14249) and F64L-S65T-GFP (sequence
15 disclosed in WO 97/11094) by polymerase chain reaction (PCR). The PCR reactions are performed according to standard protocols with the following primers:

5'ERK1:

TTggACACAAGCTTTggACACCCTCAggATATggCggCggCggCggCggCTCCggggggg
Cgggg (SEQ ID NO:7),

20 3'ERK1:

gTCATCTTCTCgAgTCTTTCAggCgCgCCCggggCCCTCTggCgCCCCTggCTgg
(SEQ ID NO:8),

5'F64L-S65T-GFP:

TTggACACAAGCTTTggACACggCgCgCCATgAgTAAaggAgAAgAACTTTTC
25 (SEQ ID NO:1)

3'F64L-S65T-GFP:

gTCATCTTCTCgAgTCTTACTCCTgAggTTTgTATAgTTCATCCATgCCATgT (SEQ ID NO:2)

To generate the mERK1-F64L-S65T-GFP (SEQ ID NO:56 & 57) fusion the ERK1
5 amplification product is digested with HindIII+AscI and the F64L-S65T-GFP product
with AscI+XhoI. To generate the F64L-S65T-GFP-mERK1 fusion the ERK1
amplification product is then digested with HindIII+Bsu36I and the F64L-S65T-GFP
product with Bsu36I+XhoI. The two pairs of digested PCR products are subsequently
ligated with a HindIII+XhoI digested plasmid (pZeoSV® mammalian expression
10 vector, Invitrogen, San Diego, CA, USA). The resulting fusion constructs are under
control of the SV40 promoter.

b) The human Erk1 gene (GenBank Accession number: X60188) was amplified using
PCR according to standard protocols with primers Erk1-top (SEQ ID NO:9) and Erk1-
bottom/+stop (SEQ ID NO:10). The PCR product was digested with restriction
15 enzymes EcoR1 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto;
GenBank Accession number U55763) digested with EcoR1 and BamH1. This produces
an EGFP-Erk1 fusion (SEQ ID NO:38 & 39) under the control of a CMV promoter.

The plamid containing the EGFP-Erk1 fusion was transfected into HEK293 cells
employing the FUGENE transfection reagent (Boehringer Mannheim). Prior to
20 experiments the cells were grown to 80%-90% confluency 8 well chambers in DMEM
with 10% FCS. The cells were washed in plain HAM F-12 medium (without FCS), and
then incubated for 30-60 minutes in plain HAM F-12 (without FCS) with 100 micromolar
PD98059, an inhibitor of MEK1, a kinase which activates Erk1; this step effectively
empties the nucleus of EGFP-Erk1. Just before starting the experiment, the HAM F-12 was
25 replaced with Hepes buffer following a wash with Hepes buffer. This removes the
PD98059 inhibitor; if blocking of MEK1 is still wanted (e.g. in control experiments), the
inhibitor is included in the Hepes buffer.

The experimental setup of the microscope was as described in example 1.

60 images were collected with 10 seconds between each, and with the test compound added after image number 10.

Addition of EGF (1-100 nM) caused within minutes a redistribution of EGFP-Erk1 from the cytoplasm into the nucleus (Fig. 9a,b).

- 5 The response was quantitated as described below and a dose-dependent relationship between EGF concentration and nuclear translocation of EGFP-Erk1 was found (Fig. 9c,d). Redistribution of GFP fluorescence is expressed in this example as the change in the ratio value between areas in nuclear versus cytoplasmic compartments of the cell. Each time profile is the average of nuclear to cytoplasmic ratios from six cells in each treatment.

10

EXAMPLE 4

Probes for detection of Erk2 redistribution.

Useful for monitoring signalling pathways involving MAPK, e.g. to identify compounds which modulate the activity of the pathway in living cells.

- 15 Erk2, a serine/threonine protein kinase, is closely related to Erk1 but not identical; it is a component of a signalling pathway that is activated by e.g. many growth factors.

- a) The rat Erk2 gene (GenBank Accession number: M64300) was amplified using PCR according to standard protocols with primers Erk2-top (SEQ ID NO:11) and Erk2-bottom/+stop (SEQ ID NO:13) The PCR product was digested with restriction enzymes Xho1 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank
20 Accession number U55763) digested with Xho1 and BamH1. This produces an EGFP-Erk2 fusion (SEQ ID NO:40 &41) under the control of a CMV promoter.

- b) The rat Erk2 gene (GenBank Accession number: M64300) was amplified using PCR according to standard protocols with primers (SEQ ID NO:11) Erk2-top and Erk2-bottom/-stop (SEQ ID NO:12). The PCR product was digested with restriction enzymes
25 Xho1 and BamH1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank

Accession number U55762) digested with XhoI and BamHI. This produces an Erk2-EGFP fusion (SEQ ID NO:58 &59) under the control of a CMV promoter.

The resulting plasmids were transfected into CHO cells and BHK cells. The cells were grown under standard conditions. Prior to experiments, the cells were starved in medium without serum for 48-72 hours. This led to a predominantly cytoplasmic localisation of both probes, especially in BHK cells. 10% fetal calf serum was added to the cells and the fluorescence of the cells was recorded as explained in example 3. Addition of serum caused the probes to redistribute into the nucleus within minutes of addition of serum.

10 EXAMPLE 5

Probes for detection of Smad2 redistribution.

Useful for monitoring signalling pathways activated by some members of the transforming growth factor-beta family, e.g. to identify compounds which modulate the activity of the pathway in living cells.

15 Smad 2, a signal transducer, is a component of a signalling pathway that is induced by some members of the TGFbeta family of cytokines.

a) The human Smad2 gene (GenBank Accession number: AF027964) was amplified using PCR according to standard protocols with primers Smad2-top (SEQ ID NO:24) and Smad2-bottom/+stop (SEQ ID NO:26) . The PCR product was digested with restriction enzymes EcoRI and Acc65I, and ligated into pEGFP-C1 (Clontech; Palo Alto; 20 GenBank Accession number U55763) digested with EcoRI and Acc65I. This produces an EGFP-Smad2 fusion (SEQ ID NO:50&51) under the control of a CMV promoter.

b) The human Smad2 gene (GenBank Accession number: AF027964) was amplified using PCR according to standard protocols with primers Smad2-top (SEQ ID NO:24) and Smad2-bottom/-stop (SEQ ID NO:25) . The PCR product was digested with restriction enzymes EcoRI and Acc65I, and ligated into pEGFP-N1 (Clontech, Palo Alto; 25

GenBank Accession number U55762) digested with EcoR1 and Acc65I. This produces a Smad2-EGFP fusion (SEQ ID NO:74 &75) under the control of a CMV promoter.

The plasmid containing the EGFP-Smad2 fusion was transfected into HEK293 cells, where it showed a cytoplasmic distribution. Prior to experiments the cells were grown in 8 well Nunc chambers in DMEM with 10% FCS to 80% confluence and starved overnight in HAM F-12 medium without FCS.

For experiments, the HAM F-12 medium was replaced with Hepes buffer pH 7.2.

The experimental setup of the microscope was as described in example 1.

90 images were collected with 10 seconds between each, and with the test compound added after image number 5.

After serum starvation of cells, each nucleus contains less GFP fluorescence than the surrounding cytoplasm (Fig. 10a). Addition of TGFbeta caused within minutes a redistribution of EGFP-Smad2 from the cytoplasm into the nucleus (Fig. 10b).

The redistribution of fluorescence within the treated cells was quantified simply as the fractional increase in nuclear fluorescence normalised to the starting value of GFP fluorescence in the nucleus of each unstimulated cell.

EXAMPLE 6

Probe for detection of VASP redistribution.

Useful for monitoring signalling pathways involving rearrangement of cytoskeletal elements, e.g. to identify compounds which modulate the activity of the pathway in living cells.

VASP, a phosphoprotein, is a component of cytoskeletal structures, which redistributes in response to signals that affect focal adhesions.

The human VASP gene (GenBank Accession number: Z46389) was amplified using PCR according to standard protocols with primers VASP-top (SEQ ID NO:94) and VASP-bottom/+stop (SEQ ID NO:95). The PCR product was digested with restriction enzymes Hind3 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with Hind3 and BamH1. This produces an EGFP-VASP fusion (SEQ ID NO:124 &125) under the control of a CMV promoter.

The resulting plasmid was transfected into CHO cells expressing the human insulin receptor using the calcium-phosphate transfection method. Prior to experiments, cells were grown in 8 well Nunc chambers and starved overnight in medium without FCS.

10 Experiments are performed in a microscope setup as described in example 1.

10% FCS was added to the cells and images were collected. The EGFP-VASP fusion was redistributed from a somewhat even distribution near the periphery into more localised structures, identified as focal adhesion points (Fig. 11).

A large number of further GFP fusions have been made or are in the process of being made, as apparent from the following Examples 7-22 which also suggest suitable host cells and substances for activation of the cellular signalling pathways to be monitored and analyzed.

EXAMPLE 7

20 **Probe for detection of actin redistribution.**

Useful for monitoring signalling pathways involving rearrangement or formation of actin filaments, e.g. to identify compounds which modulate the activity of pathways leading to cytoskeletal rearrangements in living cells.

Actin is a component of cytoskeletal structures, which redistributes in response to very many cellular signals.

The actin binding domain of the human alpha-actinin gene (GenBank Accession number:

X15804) was amplified using PCR according to standard protocols with primers ABD-top (SEQ ID NO:90) and ABD-bottom/-stop (SEQ ID NO:91). The PCR product was digested with restriction enzymes Hind3 and BamH1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with Hind3 and BamH1. This
5 produced an actin-binding-domain-EGFP fusion (SEQ ID NO:128 &129) under the control of a CMV promoter.

The resulting plasmid was transfected into CHO cells expressing the human insulin receptor. Cells were stimulated with insulin that caused the actin binding domain-EGFP probe to become redistributed into morphologically distinct membrane-associated
10 structures.

EXAMPLE 8

Probes for detection of p38 redistribution.

Useful for monitoring signalling pathways responding to various cellular stress situations,
15 e.g. to identify compounds which modulate the activity of the pathway in living cells, or as a counterscreen.

p38, a serine/threonine protein kinase, is a component of a stress-induced signalling pathway which is activated by many types of cellular stress, e.g. TNFalpha, anisomycin, UV and mitomycin C.

20 a) The human p38 gene (GenBank Accession number: L35253) was amplified using PCR according to standard protocols with primers p38-top (SEQ ID NO:14) and p38-bottom/+stop (SEQ ID NO: 16). The PCR product was digested with restriction enzymes Xho1 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with Xho1 and BamH1. This produced an EGFP-
25 p38 fusion (SEQ ID NO:46 & 47) under the control of a CMV promoter.

b) The human p38 gene (GenBank Accession number: L35253) was amplified using PCR according to standard protocols with primers p38-top (SEQ ID NO:13) and p38-

bottom/-stop (SEQ ID NO:15) . The PCR product was digested with restriction enzymes XhoI and BamHI, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with XhoI and BamHI. This produced a p38-EGFP fusion (SEQ ID NO:64 & 65) under the control of a CMV promoter.

- 5 The resulting plasmids are transfected into a suitable cell line, e.g. HEK293, in which the EGFP-p38 probe and/or the p38-EGFP probe should change its cellular distribution from predominantly cytoplasmic to nuclear within minutes in response to activation of the signalling pathway with e.g. anisomycin.

10 **EXAMPLE 9**

Probes for detection of Jnk1 redistribution.

Useful for monitoring signalling pathways responding to various cellular stress situations, e.g. to identify compounds which modulate the activity of the pathway in living cells, or as a counterscreen.

- 15 Jnk1, a serine/threonine protein kinase, is a component of a stress-induced signalling pathway different from the p38 described above, though it also is activated by many types of cellular stress, e.g. TNFalpha, anisomycin and UV.
- a) The human Jnk1 gene (GenBank Accession number: L26318) was amplified using PCR according to standard protocols with primers Jnk-top (SEQ ID NO:17) and Jnk-
- 20 bottom/+stop (SEQ ID NO:19) . The PCR product was digested with restriction enzymes XhoI and BamHI, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with XhoI and BamHI. This produced an EGFP-Jnk1 fusion (SEQ ID NO:44 &45) under the control of a CMV promoter.
- b) The human Jnk1 gene (GenBank Accession number: L26318) was amplified using PCR
- 25 according to standard protocols with primers Jnk-top (SEQ ID NO:17) and Jnk-bottom/-stop (SEQ ID NO:18) . The PCR product was digested with restriction enzymes XhoI and BamHI, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank

Accession number U55762) digested with XhoI and BamHI. This produced a Jnk1-EGFP fusion (SEQ ID NO:62 & 63) under the control of a CMV promoter.

The resulting plasmids are transfected into a suitable cell line, e.g. HEK293, in which the EGFP-Jnk1 probe and/or the Jnk1-EGFP probe should change its cellular distribution from
5 predominantly cytoplasmic to nuclear in response to activation of the signalling pathway with e.g. anisomycin.

EXAMPLE 10

Probes for detection of PKG redistribution.

10 Useful for monitoring signalling pathways involving changes in cyclic GMP levels, e.g. to identify compounds which modulate the activity of the pathway in living cells.

PKG, a cGMP-dependent serine/threonine protein kinase, mediates the guanylyl-cyclase/cGMP signal.

a) The human PKG gene (GenBank Accession number: Y07512) is amplified using PCR
15 according to standard protocols with primers PKG-top (SEQ ID NO:81) and PKG-bottom/+stop (SEQ ID NO:83). The PCR product is digested with restriction enzymes XhoI and BamHI, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with XhoI and BamHI. This produces an EGFP-PKG fusion (SEQ ID NO:134 & 135) under the control of a CMV promoter.

20 b) The human PKG gene (GenBank Accession number: Y07512) is amplified using PCR according to standard protocols with primers PKG-top (SEQ ID NO:81) and PKG-bottom/-stop (SEQ ID NO: 82). The PCR product is digested with restriction enzymes XhoI and BamHI, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with XhoI and BamHI. This produces a PKG-
25 EGFP fusion (SEQ ID NO:136 & 137) under the control of a CMV promoter.

The resulting plasmids are transfected into a suitable cell line, e.g. A10, in which the EGFP-PKG probe and/or the PKG-EGFP probe should change its cellular distribution

from cytoplasmic to one associated with cytoskeletal elements within minutes in response to treatment with agents which raise nitric oxide (NO) levels.

EXAMPLE 11

5 **Probes for detection of IkappaB kinase redistribution.**

Useful for monitoring signalling pathways leading to NFkappaB activation, e.g. to identify compounds which modulate the activity of the pathway in living cells.

IkappaB kinase, a serine/threonine kinase, is a component of a signalling pathway which is activated by a variety of inducers including cytokines, lymphokines, growth factors and
10 stress.

a) The alpha subunit of the human IkappaB kinase gene (GenBank Accession number: AF009225) is amplified using PCR according to standard protocols with primers IKK-top (SEQ ID NO:96) and IKK-bottom/+stop (SEQ ID NO:98). The PCR product is digested with restriction enzymes EcoR1 and Acc65I, and ligated into pEGFP-C1
15 (Clontech, Palo Alto; GenBank Accession number U55763) digested with EcoR1 and Acc65I. This produces an EGFP-IkappaB-kinase fusion (SEQ ID NO:120 &121) under the control of a CMV promoter.

b) The alpha subunit of the human IkappaB kinase gene (GenBank Accession number: AF009225) is amplified using PCR according to standard protocols with primers IKK-top (SEQ ID NO:96) and IKK-bottom/-stop (SEQ ID NO:97). The PCR product is
20 digested with restriction enzymes EcoR1 and Acc65I, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with EcoR1 and Acc65I. This produces an IkappaB-kinase-EGFP fusion (SEQ ID NO:122 &123) under the control of a CMV promoter.

25 The resulting plasmids are transfected into a suitable cell line, e.g. Jurkat, in which the EGFP-IkappaB-kinase probe and/or the IkappaB-kinase-EGFP probe should achieve a more cytoplasmic distribution within seconds following stimulation with e.g. TNFalpha.

EXAMPLE 12

Probes for detection of CDK2 redistribution.

Useful for monitoring signalling pathways of the cell cycle, e.g. to identify compounds
5 that modulate the activity of the pathway in living cells.

CDK2, a cyclin-dependent serine/threonine kinase, is a component of the signalling system that regulates the cell cycle.

- 10 a) The human CDK2 gene (GenBank Accession number: X61622) is amplified using PCR according to standard protocols with primers CDK2-top (SEQ ID NO:102) and CDK2-bottom/+stop (SEQ ID NO: 104). The PCR product is digested with restriction enzymes Xho1 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with Xho1 and BamH1. This produces an EGFP-CDK2 fusion (SEQ ID NO:114 &115) under the control of a CMV promoter.
- 15 b) The human CDK2 gene (GenBank Accession number: X61622) is amplified using PCR according to standard protocols with primers CDK2-top (SEQ ID NO:102) and CDK2-bottom/-stop (SEQ ID NO:103). The PCR product is digested with restriction enzymes Xho1 and BamH1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with Xho1 and BamH1. This produces a CDK2-EGFP fusion (SEQ ID NO:112 &113) under the control of a CMV promoter.
- 20 The resulting plasmids are transfected into a suitable cell line, e.g. HEK293 in which the EGFP-CDK2 probe and/or the CDK2-EGFP probe should change its cellular distribution from cytoplasmic in contact-inhibited cells, to nuclear location in response to activation with a number of growth factors, e.g. IGF.

25 EXAMPLE 13

Probes for detection of Grk5 redistribution.

Useful for monitoring signalling pathways involving desensitisation of G-protein coupled receptors, e.g. to identify compounds which modulate the activity of the pathway in living cells.

Grk5, a G-protein coupled receptor kinase, is a component of signalling pathways
5 involving membrane bound G-protein coupled receptors.

a) The human Grk5 gene (GenBank Accession number: L15388) is amplified using PCR according to standard protocols with primers Grk5-top (SEQ ID NO:27) and Grk5-bottom/+stop (SEQ ID NO:29). The PCR product is digested with restriction enzymes
10 EcoR1 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with EcoR1 and BamH1. This produces an EGFP-Grk5 fusion (SEQ ID NO:42 &43) under the control of a CMV promoter.

b) The human Grk5 gene (GenBank Accession number: L15388) is amplified using PCR according to standard protocols with primers Grk5-top (SEQ ID NO:27) and Grk5-bottom/-stop (SEQ ID NO:28). The PCR product is digested with restriction enzymes
15 EcoR1 and BamH1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with EcoR1 and BamH1. This produces a Grk5-EGFP fusion (SEQ ID NO:60 &61) under the control of a CMV promoter.

The resulting plasmids are transfected into a suitable cell line, e.g. HEK293 expressing a
20 rat dopamine D1A receptor, in which the EGFP-Grk5 probe and/or the Grk5-EGFP probe should change its cellular distribution from predominantly cytoplasmic to peripheral in response to activation of the signalling pathway with e.g. dopamine.

EXAMPLE 14

25 **Probes for detection of Zap70 redistribution.**

Useful for monitoring signalling pathways involving the T cell receptor, e.g. to identify compounds which modulate the activity of the pathway in living cells.

Zap70, a tyrosine kinase, is a component of a signalling pathway which is active in e.g. T-cell differentiation.

- a) The human Zap70 gene (GenBank Accession number: L05148) is amplified using PCR according to standard protocols with primers Zap70-top (SEQ ID NO:105) and Zap70-bottom/+stop (SEQ ID NO:107). The PCR product is digested with restriction enzymes EcoR1 and BamH1, and ligated into pEGFP-C1 (GenBank Accession number U55763) digested with EcoR1 and BamH1. This produces an EGFP-Zap70 fusion (SEQ ID NO:108 & 109) under the control of a CMV promoter.
- b) The human Zap70 gene (GenBank Accession number: L05148) is amplified using PCR according to standard protocols with primers Zap70-top (SEQ ID NO:105) and Zap70-bottom/-stop (SEQ ID NO:106). The PCR product is digested with restriction enzymes EcoR1 and BamH1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with EcoR1 and BamH1. This produces a Zap70-EGFP fusion (SEQ ID NO:110 & 111) under the control of a CMV promoter.
- The resulting plasmids are transfected into a suitable cell line, e.g. Jurkat, in which the EGFP-Zap70 probe and/or the Zap70-EGFP probe should change its cellular distribution from cytoplasmic to membrane-associated within seconds in response to activation of the T cell receptor signalling pathway with e.g. antibodies to CD3epsilon.

EXAMPLE 15

Probes for detection of p85 redistribution.

Useful for monitoring signalling pathways involving PI-3 kinase, e.g. to identify compounds which modulate the activity of the pathway in living cells.

p85alpha is the regulatory subunit of PI3-kinase which is a component of many pathways involving membrane-bound tyrosine kinase receptors and G-protein-coupled receptors.

- a) The human p85alpha gene (GenBank Accession number: M61906) was amplified using PCR according to standard protocols with primers p85-top-C (SEQ ID NO:22) and p85-

bottom/+stop (SEQ ID NO:23) . The PCR product was digested with restriction enzymes Bgl2 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with Bgl2 and BamH1. This produced an EGFP-p85alpha fusion (SEQ ID NO:48 &49) under the control of a CMV promoter.

- 5 b) The human p85alpha gene (GenBank Accession number: M61906) was amplified using PCR according to standard protocols with primers p85-top-N (SEQ ID NO:20) and p85-bottom/-stop (SEQ ID NO:21) . The PCR product was digested with restriction enzymes EcoR1 and BamH1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with EcoR1 and BamH1. This produced
10 a p85alpha-EGFP fusion (SEQ ID NO:66 &67) under the control of a CMV promoter.

The resulting plasmids are transfected into a suitable cell line, e.g. CHO expressing the human insulin receptor, in which the EGFP-p85 probe and/or the p85-EGFP probe may change its cellular distribution from cytoplasmic to membrane-associated within minutes in response to activation of the receptor with insulin.

15

EXAMPLE 16

Probes for detection of protein-tyrosine phosphatase redistribution.

Useful for monitoring signalling pathways involving tyrosine kinases, e.g. to identify compounds which modulate the activity of the pathway in living cells.

- 20 Protein-tyrosine phosphatase1C, a tyrosine-specific phosphatase, is an inhibitory component in signalling pathways involving e.g. some growth factors.

- a) The human protein-tyrosine phosphatase 1C gene (GenBank Accession number: X62055) is amplified using PCR according to standard protocols with primers PTP-top (SEQ ID NO:99) and PTP-bottom/+stop (SEQ ID NO:101). The PCR product is
25 digested with restriction enzymes Xho1 and EcoR1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with Xho1 and EcoR1. This produces an EGFP-PTP fusion (SEQ ID NO:116 & 117) under the control

of a CMV promoter.

- b) The human protein-tyrosine phosphatase 1C gene (GenBank Accession number: X62055) is amplified using PCR according to standard protocols with primers PTP-top (SEQ ID NO:99) and PTP-bottom/-stop (SEQ ID NO:100). The PCR product is
- 5 digested with restriction enzymes Xho1 and EcoR1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with Xho1 and EcoR1. This produces a PTP-EGFP fusion (SEQ ID NO:118 & 119) under the control of a CMV promoter.

The resulting plasmids are transfected into a suitable cell line, e.g. MCF-7 in which the

10 EGFP-PTP probe and/or the PTP-EGFP probe should change its cellular distribution from cytoplasm to the plasma membrane within minutes in response to activation of the growth inhibitory signalling pathway with e.g. somatostatin.

EXAMPLE 17

15 **Probes for detection of Smad4 redistribution.**

Useful for monitoring signalling pathways involving most members of the transforming growth factor-beta family, e.g. to identify compounds which modulate the activity of the pathway in living cells.

Smad4, a signal transducer, is a common component of signalling pathways induced by

20 various members of the TGFbeta family of cytokines.

- a) The human Smad4 gene (GenBank Accession number: U44378) was amplified using PCR according to standard protocols with primers Smad4-top and Smad4-bottom/+stop (SEQ ID NO:35) . The PCR product was digested with restriction enzymes EcoR1 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number
- 25 U55763) digested with EcoR1 and BamH1. This produce an EGFP-Smad4 fusion (SEQ ID NO:52 & 53) under the control of a CMV promoter.

- b) The human Smad4 gene (GenBank Accession number: U44378) was amplified using

PCR according to standard protocols with primers Smad4-top (SEQ ID NO:33) and Smad4-bottom/-stop (SEQ ID NO:34). The PCR product was digested with restriction enzymes EcoR1 and BamH1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with EcoR1 and BamH1. This produced
5 a Smad4-EGFP fusion (SEQ ID NO:76 & 77) under the control of a CMV promoter.

The resulting plasmids are transfected into a cell line, e.g. HEK293 in which the EGFP-Smad4 probe and/or the Smad4-EGFP probe should change its cellular distribution within minutes from cytoplasmic to nuclear in response to activation of the signalling pathway with e.g. TGFbeta.

10

EXAMPLE 18

Probes for detection of Stat5 redistribution.

Useful for monitoring signalling pathways involving the activation of tyrosine kinases of the Jak family, e.g. to identify compounds that modulate the activity of the pathway in
15 living cells.

Stat5, signal transducer and activator of transcription, is a component of signalling pathways that are induced by e.g. many cytokines and growth factors.

- a) The human Stat5 gene (GenBank Accession number: L41142) was amplified using
20 PCR according to standard protocols with primers Stat5-top (SEQ ID NO:30) and Stat5-bottom/+stop (SEQ ID NO:32). The PCR product was digested with restriction enzymes Bgl2 and Acc65I, and ligated into pEGFP-C1 (Clontech; Palo Alto; GenBank Accession number U55763) digested with Bgl2 and Acc65I. This produced an EGFP-Stat5 fusion (SEQ ID NO:54 & 55) under the control of a CMV promoter.
- 25 b) The human Stat5 gene (GenBank Accession number: L41142) was amplified using PCR according to standard protocols with primers Stat5-top (SEQ ID NO:30) and Stat5-bottom/-stop (SEQ ID NO:331). The PCR product was digested with restriction

enzymes Bgl2 and Acc65I, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with Bgl2 and Acc65I. This produced a Stat5-EGFP fusion (SEQ ID NO:78 & 79) under the control of a CMV promoter.

- 5 The resulting plasmids are transfected into a suitable cell line, e.g. MIN6 in which the EGFP-Stat5 probe and/or the Stat5-EGFP probe should change its cellular distribution from cytoplasmic to nuclear within minutes in response to activation signalling pathway with e.g. prolactin.

EXAMPLE 19

10 **Probes for detection of NFAT redistribution.**

Useful for monitoring signalling pathways involving activation of NFAT, e.g. to identify compounds which modulate the activity of the pathway in living cells.

NFAT, an activator of transcription, is a component of signalling pathways involved in e.g. immune responses.

- 15 a) The human NFAT1 gene (GenBank Accession number: U43342) is amplified using PCR according to standard protocols with primers NFAT-top (SEQ ID NO:84) and NFAT-bottom/+stop (SEQ ID NO:86). The PCR product is digested with restriction enzymes Xho1 and EcoR1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with Xho1 and EcoR1. This produces an EGFP-
20 NFAT fusion (SEQ ID NO:130 & 131) under the control of a CMV promoter.

- b) The human NFAT gene (GenBank Accession number: U43342) is amplified using PCR according to standard protocols with primers NFAT-top (SEQ ID NO:84) and NFAT-bottom/-stop (SEQ ID NO:85). The PCR product is digested with restriction enzymes Xho1 and EcoR1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank
25 Accession number U55762) digested with Xho1 and EcoR1. This produces an NFAT-EGFP fusion (SEQ ID NO:132 & 133) under the control of a CMV promoter.

The resulting plasmids are transfected into a suitable cell line, e.g. Jurkat, in which the

EGFP-NFAT probe and/or the NFAT-EGFP probe should change its cellular distribution from cytoplasmic to nuclear within minutes in response to activation of the signalling pathway with e.g. antibodies to CD3epsilon.

5 **EXAMPLE 20**

Probes for detection of NFkappaB redistribution.

Useful for monitoring signalling pathways leading to activation of NFkappaB, e.g. to identify compounds which modulate the activity of the pathway in living cells.

10 NFkappaB, an activator of transcription, is a component of signalling pathways that are responsive to a variety of inducers including cytokines, lymphokines, and some immunosuppressive agents.

a) The human NFkappaB p65 subunit gene (GenBank Accession number: M62399) is amplified using PCR according to standard protocols with primers NFkappaB-top (SEQ ID NO:87) and NFkappaB-bottom/+stop (SEQ ID NO:89). The PCR product is
15 digested with restriction enzymes XhoI and BamHI, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with XhoI and BamHI. This produces an EGFP-NFkappaB fusion (SEQ ID NO:142 & 143) under the control of a CMV promoter.

b) The human NFkappaB p65 subunit gene (GenBank Accession number: M62399) is
20 amplified using PCR according to standard protocols with primers NFkappaB-top (SEQ ID NO:87) and NFkappaB-bottom/-stop (SEQ ID NO:88). The PCR product is digested with restriction enzymes XhoI and BamHI, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with XhoI and BamHI. This produces an NFkappaB-EGFP fusion (SEQ ID NO:140 & 141) under the control of a
25 CMV promoter.

The resulting plasmids are transfected into a suitable cell line, e.g. Jurkat, in which the EGFP-NFkappaB probe and/or the NFkappaB-EGFP probe should change its cellular

distribution from cytoplasmic to nuclear in response to activation of the signalling pathway with e.g. TNFalpha.

EXAMPLE 21

5 **Probe for detection of RhoA redistribution.**

Useful for monitoring signalling pathways involving RhoA, e.g. to identify compounds which modulate the activity of the pathway in living cells.

RhoA, a small GTPase, is a component of many signalling pathways, e.g. LPA induced cytoskeletal rearrangements.

- 10 The human RhoA gene (GenBank Accession number: L25080) was amplified using PCR according to standard protocols with primers RhoA-top (SEQ ID NO:92) and RhoA-bottom/+stop (SEQ ID NO:93). The PCR product was digested with restriction enzymes Hind3 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with Hind3 and BamH1. This produced an EGFP-RhoA fusion
15 (SEQ ID NO:126 & 127) under the control of a CMV promoter.

The resulting plasmid is transfected into a suitable cell line, e.g. Swiss3T3, in which the EGFP-RhoA probe should change its cellular distribution from a reasonably homogenous to a peripheral distribution within minutes of activation of the signalling pathway with e.g. LPA.

20

EXAMPLE 22

Probes for detection of PKB redistribution.

Useful for monitoring signalling pathways involving PKB e.g. to identify compounds which modulate the activity of the pathway in living cells.

- 25 PKB, a serine/threonine kinase, is a component in various signalling pathways, many of

which are activated by growth factors.

- a) The human PKB gene (GenBank Accession number: M63167) is amplified using PCR according to standard protocols with primers PKB-top (SEQ ID NO:36) and PKB-bottom/+stop (SEQ ID NO:80). The PCR product is digested with restriction enzymes Xho1 and BamH1, and ligated into pEGFP-C1 (Clontech, Palo Alto; GenBank Accession number U55763) digested with Xho1 and BamH1. This produces an EGFP-PKB fusion (SEQ ID NO:138 & 139) under the control of a CMV promoter.
- b) The human PKB gene (GenBank Accession number: M63167) was amplified using PCR according to standard protocols with primers PKB-top (SEQ ID NO:36) and PKB-bottom/-stop (SEQ ID NO:37). The PCR product was digested with restriction enzymes Xho1 and BamH1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with Xho1 and BamH1. This produced a PKB-EGFP fusion (SEQ ID NO:70 & 71) under the control of a CMV promoter.
- The resulting plasmids are transfected into a suitable cell line, e.g. CHO expressing the human insulin receptor, in which the EGFP-PKB probe and/or the PKB-EGFP probe cycles between cytoplasmic and membrane locations during the activation-deactivation process following addition of insulin. The transition should be apparent within minutes.

EXAMPLE 23

Measurement of the real-time redistribution of protein kinase C α isoform-GFP fusion (PKC α -GFP) in response to carbamylcholine stimulation of the muscarinic M1 receptor; 96 parallel redistribution measurements in microtiter plates.

BHK cells were stably expressing a recombinant human muscarinic type 1 receptor, under the selection with 500 μ g/ml Methotrexate, and also a PKC α -GFP construct (K α A 048), under the selection of 500 nM Zeocin. The cells were grown in 96-well plates (Packard ViewPlate, black with transparent bottom), washed and preincubated in a Hank's Buffered

Salt solution (HBSS) without phenol red, with 20 mM HEPES and 5.5 mM glucose.

The plate was measured in a FLIPR™ (Fluorescence Imaging Plate Reader) from Molecular Devices. The 488 nm emission line from an argon ion laser, run at between 0.4 and 0.8 W output, was used to excite fluorescence from the GFP. Emission wavelengths
5 were collected through a 510 to 565 nm band pass filter.

The cells were challenged with three doses of carbamylcholine, an M1 receptor agonist known from previous studies to give a microscopically detectable redistribution of the PKC α -GFP construct [(Almholt *et al.* 1997)]. Measurements were made every 10 seconds for 5 minutes. After data handling including normalisation of baseline fluorescence for the
10 different wells, background subtraction and averaging the 6 wells used for each concentration the data presented in figure 14 were obtained. It can clearly be seen (Fig 14) that carbamylcholine gave a time- and dose-dependent, and transient, decrease in fluorescence very similar to the time- and dose-dependent profile seen in microscopic fluorescence measurements [(see Almholt *et al.* 1997)]. This experiment was repeated
15 twice on the same batch of cells with similar results.

EXAMPLE 24

**Measurement of the real-time redistribution of cyclic-AMP dependent protein kinase catalytic subunit-GFP fusion (C-GFP^{LT}) in response to forskolin stimulation of the
20 adenylate cyclase; 96 parallel redistribution measurements in microtiter plates.**

CHO cells were stably transfected with hybrid DNA for the PKA catalytic subunit-F64L+S65T GFP (C-GFP^{LT}) fusion protein, and were typically under continuous selection with 1000 μ g/ml zeocin (Invitrogen). The cells were grown without selection for 2 days in 96-well plates (Packard ViewPlate, black with transparent bottom), washed and
25 preincubated in a Hank's Buffered Salt solution (HBSS) without phenol red, with 20 mM HEPES and 5.5 mM glucose.

The plate was measured in a FLIPR™ (Fluorescence Imaging Plate Reader) from Molecular Devices. The 488 nm emission line from an argon ion laser, run at between 0.4

and 0.8 W output, was used to excite fluorescence from the GFP. Emission wavelengths were collected through a 510 to 565 nm band pass filter.

The cells were challenged with three doses of forskolin (Fig 15), an adenylate cyclase agonist known from previous studies to give a microscopically detectable redistribution of the C-GFP^{LT} construct [(Almholt *et al.* 1998)]. Measurements were made every 10 seconds for over 6 minutes from the point of addition of forskolin. After data handling including normalisation of baseline fluorescence for the different wells, background subtraction and averaging the 6 wells used for each concentration the data presented below were obtained. It can clearly be seen in figure 15 that forskolin gave a time- and dose-dependent decrease in fluorescence very similar to the time- and dose-dependent profile seen in microscopic fluorescence measurements [(see Almholt *et al.* 1998)]. This experiment was repeated twice on the same batch of cells with similar results.

EXAMPLE 25

Measurement of the redistribution response of cyclic-AMP dependent protein kinase catalytic subunit-GFP fusion (C-GFP^{LT}) after forskolin stimulation of the adenylate cyclase; measurement of the change in total fluorescence upon permeabilisation of agonist-treated cells.

CHO cells were stably transfected with hybrid DNA for the PKA catalytic subunit-F64L+S65T GFP (C-GFP^{LT}) fusion protein, and were typically under continuous selection with 1000 µg/ml zeocin (Invitrogen). For the experiments reported here, cells were grown without selection to 90% confluence in 8-well tissue culture-treated Lab-Tek® chambered coverglass units (chambers, obtained from Nunc, Inc. Illinois, USA). Immediately prior to the experiment growth medium was washed from the cells and replaced with 200 µl HEPES buffer per well.

For the results reported here, chambers were measured using a cooled CCD camera (KAF1400 chip, Photometrics Ltd., USA) attached to an inverted microscope (Diaphot 300, Nikon, Japan) equipped with a x40 oil-immersion Fluor lens, NA 1.4. Cells were

illuminated with 450-490 nm light from a 50 W HBO lamp, and emitted light collected between 510-560 nm.

The cells were challenged with four doses of forskolin, an adenylate cyclase agonist known from previous studies to give a microscopically detectable redistribution of the C-GFP^{LT} construct [(Almholt *et al.* 1998)]. Images were collected at 10-second intervals for a period of 10 minutes for each treatment. Six minutes after the addition of forskolin or buffer control, Triton-X100 was added to a final concentration of 0.1%. The detergent releases freely mobile C-GFP^{LT} from the cells. The change in fluorescence resulting from this loss was measured after 1 minute of equilibration. After data handling including background subtraction and normalisation to pre-detergent values, the data presented in figure 16 were obtained. Permeabilisation caused decreases in fluorescence, the magnitude of which were dependent on the forskolin treatments. The dose-dependent profile for forskolin activation of the cAMP system as revealed by this method was very similar to that registered by other methods (see Almholt *et al.* 1998). This experiment was repeated twice on the same batch of cells with similar results.

EXAMPLE 26

Probe for detection of PKCbeta2 redistribution.

Useful for monitoring signalling pathways involving protein kinase C, e.g. for identifying compounds which modulate the activity of the pathway in living cells.

PKCbeta2, a serine/threonine protein kinase, is closely related to PKCalpha but not identical; it is a component of a signalling pathway that is activated by elevation of intracellular calcium concomitant with an increase in diacylglycerol species.

a) The human PKCbeta2 gene (GenBank Accession number: X07109) was amplified using PCR according to standard protocols with primers PKCbeta2-top (SEQ ID NO:162) and PKCbeta2-bottom (SEQ ID NO:163). The PCR product was digested with restriction enzymes XhoI and BamHI, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with XhoI and BamHI. This produces a PKCbeta2-

EGFP fusion (SEQ ID NO:146 & 147) under the control of a CMV promoter.

The resulting plasmids are transfected into BHK cells transfected with a human muscarinic acetylcholine receptor type M1. The cells are grown under standard conditions. The fluorescence of the cells is recorded as explained in example 3. Addition of 1 μ M -100 μ M carbachol causes a transient redistribution of fluorescence within the cells whereby it changes from a cytosolic location to the plasma membrane.

EXAMPLE 27

Probes for detection of PDE4D redistribution.

- 10 Useful for monitoring signalling pathways involving Protein Kinase A, e.g. to identify compounds which modulate the activity of the pathway in living cells.

PDE4D3, PDE4D4 and PDE4D5 are closely related splicing variants of PDE4D, a cAMP dependent phosphodiesterase. They are components of signalling pathways which involves cAMP.

- 15 The human PDE4D3, PDE4D4 and PDE4D5 genes (GenBank Accession numbers: L20970, L20969 and AF012073) are amplified using PCR according to standard protocols with the common bottom primer PDE4D-bottom (SEQ ID NO:159) and PDE4D3-top (SEQ ID NO:156), PDE4D4-top (SEQ ID NO:157) and PDE4D5-top respectively (SEQ ID NO:158) The PCR products are digested with restriction enzymes Hind3 and EcoR1, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with Hind3 and EcoR1. This produces a PDE4D3-EGFP fusion (SEQ ID NO:154 & 155), a PDE4D4-EGFP fusion (SEQ ID NO:150 & 151) and a PDE4D5-EGFP fusion (SEQ ID NO:148 & 149), all three under the control of a CMV promoter.

- 25 The resulting plasmids are transfected into MVLEC cells. The cells are grown under standard conditions. The fluorescence of the cells is recorded as explained in example 3. Addition of test compounds may cause a redistribution of fluorescence within the cells from an organised cytosolic distribution to a more random one.

EXAMPLE 28

Probes for detection of PDE5 redistribution.

Useful for monitoring signalling pathways involving Protein Kinase G, e.g. to identify
5 compounds which modulate the activity of the pathway in living cells.

PDE5 is a cGMP specific phosphodiesterase. It is a component of a signalling pathway which is activated by e.g. nitric oxide.

a) The human PDE5 gene (GenBank Accession numbers: AJ004865) is amplified using PCR according to standard protocols with primers PDE5-top (SEQ ID NO:160) and PDE5-
10 bottom (SEQ ID NO:161). The PCR product is digested with restriction enzymes EcoR1 and Acc65I, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number U55762) digested with EcoR1 and Acc65I. This produces a PDE5-EGFP fusion (SEQ ID NO 144 & 145) under the control of a CMV promoter.

The resulting plasmids are transfected into e.g. A10 cells. The cells are grown under
15 standard conditions. The fluorescence of the cells is recorded as explained in example 3. Addition of test compounds may cause a redistribution of fluorescence within the cells from an organized cytosolic distribution to a more random one.

EXAMPLE 29

Probe for detection of Ikappa-kinase redistribution.

The human IKKbeta (GenBank Acc. No. AF031416) is amplified using PCR according to standard protocols with primers IKKbeta-top (SEQ ID NO:164) and IKKbeta-bottom (SEQ ID NO:165). The PCR product is digested with restriction enzymes Hind3 and Acc65I, and ligated into pEGFP-N1 (Clontech, Palo Alto; GenBank Accession number
25 U55762) digested with Hind3 and Acc65I. This produces a IKKbeta-EGFP fusion (SEQ ID NO 152 & 153) under the control of a CMV promoter.

EXAMPLE 30

Construction of catalytically inactive Erk1 probes.

A catalytically inactive probe has the advantage that it interferes less with the normal
5 physiology of the cell while retaining its ability to report on activation of a cellular
signalling pathway by redistribution.

The Erk1 probes described above in Example 3 were subjected to site specific mutagenesis
which specifically replaced the lysine at amino acid residue number 71 in the native Erk1
sequence with arginine. This mutation is known to inactivate the catalytic activity of Erk1.
10 The redistribution patterns of the inactive Erk1 probes were identical to the original Erk1
probes, i.e. they reported on activation of the pathway by redistributing from the cytoplasm
into the nucleus. The establishment of stable cell lines expressing the probe was facilitated.

REFERENCES:

- Adams, S.R., Harootunian, A.T., Buechler, Y.J., Taylor, S.S. & Tsien, R.Y. (1991) *Nature* **349**, 694-697.
- 5 Almholt, K., Arkhammar, P.O.G., Thastrup, O., & Tullin, S. (1997) *Mol. Biology of the Cell* **8**: Suppl. S: 72.
- Almholt, K., Terry, B.R., Skyggebjerg, O., Scudder, K., Tullin, S., Thastrup, O. (1998) [Manuscript], see Appendix II.
- Barak, L.S., Ferguson, S.S.G., Zhang, J. & Caron, M.G. (1997) *J. Biol. Chem.* **272**:44,
10 27497-27500.
- Bastiaens, P.I.H. & Jovin, T.M. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 8407-8412.
- Beals, C.R., Clipstone, N.A., Ho, S.N. & Crabtree, G.R. (1998) *Genes and Development* **11**:7, 824-834.
- Blobe, G.C., Stribling, D.S., Fabbro, D., Stabel, S & Hannun, Y.A. (1996) *J. Biol. Chem.*
15 **271**, 15823-15830.
- Carey, K.L., Richards, S.A., Lounsbury, K.M. & Macara, I.G. (1996) *J. Cell Biol.* **133**: 5, 985-996.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. & Prasher, D.C. (1994) *Science* **263**, 802-805.
- 20 Cossette, L.J., Hoglinger, O., Mou, L.J. & Shen, S.H. (1997) *Exp. Cell Res.* **223**, 459-466.
- DeBernardi, M.A. & Brooker, G. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 4577-4582.
- Farese, R.V. (1992) *Biochem. J.* **288**, 319-323.
- Fulop Jr., T., Leblanc, C., Lacombe, G. & Dupuis, G. (1995) *FEBS Lett.* **375**, 69-74.

- Georget, V., Lobaccaro, J.M., Terouanne, B., Mangeat, P., Nicolas, J.C. & Sultan, C. (1997) *Mol. Cell. Endocrinol.* **129**:1, 17-26.
- Godson, C., Masliah, E., Balboa, M.A., Ellisman, M.H. & Insel, P.A. (1996) *Biochem. Biophys. Acta* **1313**, 63-71.
- 5 Guiliano, K.A., DeBiasio, R., Dunlay, R.T., Gough, A., Volosky, J.M., Zock, J., Pavlakis, G.N. & Taylor, D.L. (1997) *J. Biomol. Screening* **2**:4, 249-259.
- Khalil, R.A., Lajoie, C., Resnick, M.S. & Morgan, K.G. (1992) *Am. J. Physiol.* **263** (Cell Physiol. 32) C714-C719.
- Oancea, E., Teruel, M.N., Quest, A.F.G. & Meyer, T. (1998) *J. Cell Biol.* **140**:3, 485-498.
- 10 Sano, M., Kohno, M. & Iwanaga, M. (1995) *Brain Res.* **688**, 213-218.
- Sakai, N., Sasaki, K., Hasegawa, C., Ohkura, M., Sumioka, K., Shirai, Y. & Saito, N. (1996) *Soc. Neuroscience* **22**, 69P (Abstract).
- Sakai, N., Sakai, K., Hasegawa, C., Ohkura, M., Sumioka, K., Shirai, Y., & Naoaki, S. (1997) *Japanese Journal of Pharmacology* **73**, 69P (Abstract of a meeting held 22-23 March).
- 15 Schmidt, D.J., Ikebe, M., Kitamura, K. & Fay, F.S. (1997) *FASEB J.* **11**, 2924 (Abstract).
- Schroeder, K. & Neagle, B.J. (1996) *Biomolecular Screening* **1**, 75-80.
- Silverman, L., Campbell, R. & Broach, J.R. (1998) *Current Opinion in Chemical Biology* **2**:3, 397-403.
- 20 Stauffer, T.P., Ahn, S. & Meyer, T. (1998) *Current Biol.* **8**:6, 343-346.

CLAIMS

1. A method for extracting quantitative information relating to an influence on a cellular response, the method comprising recording variation, caused by the influence on mechanically intact or permeabilised living cells, in spatially distributed light emitted from a luminophore, the luminophore being present in the cells and being capable of being redistributed in a manner which is related with the degree of the influence, and/or of being modulated by a component which is capable of being redistributed in a manner which is related to the degree of the influence, resulting in a modulation of the luminescence characteristics of the luminophore, and processing the recorded variation in the luminescence characteristics to provide quantitative information correlating the recorded variation to the degree of the influence on the cellular response.
2. A method according to claim 1 for extracting quantitative information relating to an influence on an intracellular pathway involving redistribution of at least one component associated with the pathway, or part thereof, the method comprising recording the result of the influence on mechanically intact or permeabilised living cells, as manifested in spatially distributed light emitted from a luminophore which is present in the cells and which is capable of being redistributed, by modulation of the pathway, in a manner which is related to the redistribution of the at least one component of the intracellular pathway, processing the recorded result to provide quantitative information correlating the change in the measured property of the light to the degree of the influence on the intracellular pathway.
3. A method according to claim 1 or 2, wherein the quantitative information which is indicative of the degree of the cellular response to the influence or the result of the influence on the intracellular pathway is extracted from the recorded variation according to a predetermined calibration based on responses or results, recorded in the same manner, to known degrees of a relevant specific influence.
4. A method according to any of claims 1-3, wherein the influence comprises contact between the mechanically intact or permeabilised living cells and a chemical substance and/or incubation of the mechanically intact or permeabilised living cells with a chemical substance.

5. A method according to any of claims 1-4, wherein the influence is a substance whose effect on an intracellular pathway is to be determined.
6. A method according to any of claims 1-5, wherein the cells comprise a group of cells contained within a spatial limitation.
- 5 7. A method according to any of claims 1-5, wherein the cells comprise multiple groups of cells contained within multiple spatial limitations.
8. A method according to any of claims 1-7, wherein the cells comprise multiple groups of cells that are qualitatively the same but are subjected to different influences.
9. A method according to any of claims 1-7, wherein the cells comprise multiple groups
10 of cells that are qualitatively different but are subjected to the same influence.
10. A method according to any of claims 1-9, wherein the recording is performed by means of a detector capable of measuring total luminescence in a non-spatially resolved fashion, the recording comprising a time series of measurements of the total luminescence of the cells of one or several of the spatial limitations.
- 15 11. A method according to claim 10, wherein the signal is measured from individual spatial limitations one at a time, the recording being made in the individual spatial limitation by means of an apparatus to sequentially position each one of the limitations in the field of view of the detector, and repeating the positioning and measuring process until all of the spatial limitations have been measured.
- 20 12. A method according to claim 11, wherein the detector is a photomultiplier tube (PMT).
13. A method according to any of claims 1-9, wherein more than one of the spatial limitations are measured simultaneously.
14. A method according to claim 13, wherein the multiple spatial limitations are measured
25 simultaneously by means of a one- or two-dimensional array detector, whereby the multiple spatial limitations are imaged onto the array detector such that discrete subsets of the detecting units (pixels) in the array detector measure the signal from one and

only one of the multiple spatial limitations, the signal from any one spatial limitation being the combined signal from those pixels that receive the image from one of the spatial limitations.

15. A method according to claim 14, wherein the detector is a linear diode array.
- 5 16. A method according to claim 14, wherein the detector is a video camera.
17. A method according to claim 14, wherein the detector is a charge transfer device.
18. A method according to claim 17, wherein the charge transfer device is a charge-coupled device.
- 10 19. A method according to any of claims 1-18, wherein the luminophore must be illuminated in order to emit light.
20. A method according to any of claims 13-18, wherein all of the multiple spatial limitations are simultaneously illuminated during the measurement operation.
21. A method according to any of claims 10-18, wherein the individual spatial limitations are singly illuminated only during the time period in which they are being measured.
- 15 22. A method according to any of claims 10-18, wherein the illumination is provided by a laser which is scanned in a raster fashion over some or all of the spatial limitations being measured, the scanning taking place at a rate substantially faster than the measurement process such that the illumination appears to the measurement process to be continuous in time and spatially uniform over the region being measured.
- 20 23. A method according to any of claims 1-22, wherein the spatial limitations are spatial limitations arranged in one or more arrays on a common carrier.
24. A method according to claim 23, wherein the spatial limitations are wells in a plate of microtiter type.
- 25 25. A method according to any of claims 1-22 wherein the spatial limitations are domains defined on a substrate on which the cells are present.

26. A method according to claim 25 wherein the domains are domains established by the presence of the cells on the substrate in a pattern defining the domains.
27. A method according to claim 25 wherein the domains are domains established by the spatial pattern of the influence as it is applied to or contacted with the cells.
- 5 28. A method according to any of claims 1-27, wherein the recording is performed at a series of points in time, in which the application of the influence occurs at some time after the first time point in the series of recordings, the recording being performed, e.g., with a predetermined time spacing of from 0.1 seconds to 1 hour, preferably from 1 to 60 seconds, more preferably from 1 to 30 seconds, in particular from 1 to 10 seconds, 10 over a time span of from 1 second to 12 hours, such as from 10 seconds to 12 hours, e.g., from 10 seconds to one hour, such as from 60 seconds to 30 minutes or 20 minutes.
29. A method according to claim 28, wherein the recording is made at two points in time, one point being before, and the other point being after the application of the influence.
- 15 30. A method according to any of claims 1-29, wherein the cells are fixed at a point in time after the application of the influence at which the response has been predetermined to be significant, and the recording is made at an arbitrary later time.
31. A method according to any of claims 1-30, wherein the luminophore is a luminophore that is capable of being redistributed in a manner that is physiologically relevant to the 20 degree of the influence.
32. A method according to any of claims 1-30, wherein the luminophore is a luminophore which is capable of associating with a component which is capable of being redistributed in manner which is physiologically relevant to the degree of the influence.
33. A method according to any of claims 1-30, wherein the luminophore is a luminophore 25 which is capable of being redistributed in a manner which is experimentally determined to be correlated to the degree of the influence.
34. A method according to any of claims 1-30, wherein the luminophore is a luminophore

which is capable of being redistributed, by modulation of the intracellular pathway, in substantially the same manner as the at least one component of the intracellular pathway.

35. A method according to any of claims 1-30, wherein the luminophore is a luminophore
5 which is capable of being quenched upon spatial association with a component which is redistributed by modulation of the pathway, the quenching being measured as a decrease in the intensity of the luminescence.
36. A method according to any of claims 1-30, wherein the variation in spatially
10 distributed light emitted by the luminophore is detected by a change in the resonance energy transfer between the luminophore and another luminescent entity capable of delivering energy to the luminophore, each of which has been selected or engineered to become part of, bound to or associated with particular components of the intracellular pathway, and one of which undergoes redistribution in response to the influence,
15 thereby changing the amount of resonance energy transfer, the change in the resonance energy transfer being measured as a change in the intensity of emission from the luminophore.
37. A method according to any of claims 1-35, wherein the intensity of the light being recorded is a function of the fluorescence lifetime, polarisation, wavelength shift, or other property which is modulated as a result of the underlying cellular response.
- 20 38. A method according to any of claims 1-37, wherein the light to be measured passes through a filter which selects the desired component of the light to be measured and rejects other components.
39. A method according to any of claims 2-38, wherein the intracellular pathway is an intracellular signalling pathway.
- 25 40. A method according to any of claims 1-39, wherein the luminophore is a fluorophore.
41. A method according to any of claims 1-40, wherein the luminophore is a polypeptide encoded by and expressed from a nucleotide sequence harboured in the cells.

42. A method according to any of claims 1-41 for detecting intracellular redistribution of a biologically active polypeptide affecting intracellular processes upon activation, the method comprising
- a) culturing one or more cells containing a nucleotide sequence coding for a hybrid polypeptide comprising a GFP which is N- or C-terminally tagged, optionally through a linker, to a biologically active polypeptide under conditions permitting expression of the nucleotide sequence,
 - b) modulating the activity of the biologically active polypeptide by incubating the cells with a substance having biological activity, and
 - c) measuring the fluorescence produced by the incubated cells and determining the result or variation with respect to the fluorescence, such result or variation being indicative of the redistribution of a biologically active polypeptide in said cells.
43. A method according to claim 42, wherein the luminophore is a hybrid polypeptide comprising a fusion of at least a portion of each of two polypeptides one of which comprises a luminescent polypeptide and the other one of which comprises a biologically active polypeptide, as defined herein.
44. A method according to claim 43, wherein the luminescent polypeptide is a GFP as defined herein.
45. A method according to claim 44, wherein the GFP is selected from the group consisting of green fluorescent proteins having the F64L mutation as defined herein.
46. A method according to claim 45, wherein the GFP is a GFP variant selected from the group consisting of F64L-GFP, F64L-Y66H-GFP, F64L-S65T-GFP, and EGFP.
47. A method according to claim 42, wherein the nucleotide sequence is a DNA sequence.
48. A method according to claims 42-47, wherein the modulation is activation.
49. A method according to claims 42-47, wherein the modulation is deactivation.
50. A method according to any of claims 1-49, wherein the cells are selected from the

group consisting of fungal cells, such as yeast cells; invertebrate cells including insect cells; and vertebrate cells, such as mammalian cells.

51. A method according to claim 50, wherein the mechanically intact or permeabilised living cells are mammalian cells which, during the time period over which the influence is observed, are incubated at a temperature of 30°C or above, preferably at a temperature of from 32°C to 39°C, more preferably at a temperature of from 35°C to 38°C, and most preferably at a temperature of about 37°C.
52. A method according to any of claims 1-51, wherein the mechanically intact or permeabilised living cells are part of a matrix of identical or non-identical cells.
53. A method according to any of claims 41-52, wherein the nucleotide sequence has been introduced into the cells in the form of a nucleic acid construct coding for a fusion polypeptide comprising a biologically active polypeptide that is a component of an intracellular signalling pathway, or a part thereof, and a GFP.
54. A method according to claim 53, wherein the nucleic acid construct is a nucleic acid construct coding for a fusion polypeptide comprising a biologically active polypeptide that is a component of an intracellular signalling pathway, or a part thereof, and an F64L mutant of GFP.
55. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 53 or 54, wherein the biologically active polypeptide is a protein kinase or a phosphatase.
56. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 53 – 55, wherein the GFP is N- or C-terminally tagged, optionally via a peptide linker, to the biologically active polypeptide or part thereof.
57. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 53, 54 or 56, wherein the biologically active polypeptide is a transcription factor or a part thereof which changes cellular localisation upon activation.

58. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 53, 54 or 56, wherein the biologically active polypeptide is a protein, or a part thereof, which is associated with the cytoskeletal network and which changes cellular localisation upon activation.
- 5 59. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to any of claims 53-56, wherein the biologically active polypeptide is a protein kinase or a part thereof which changes cellular localisation upon activation.
- 10 60. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 59, wherein the protein kinase is a serine/threonine protein kinase or a part thereof capable of changing intracellular localisation upon activation.
- 15 61. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 59, wherein the protein kinase is a tyrosine protein kinase or a part thereof capable of changing intracellular localisation upon activation.
62. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 59, wherein the protein kinase is a phospholipid-dependent serine/threonine protein kinase or a part thereof capable of changing intracellular localisation upon activation.
- 20 63. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 59, wherein the protein kinase is a cAMP-dependent protein kinase or a part thereof capable of changing cellular localisation upon activation.
- 25 64. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 63 which codes for a PKAc-F64L-S65T-GFP fusion.
65. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 59, wherein the protein kinase is a cGMP-dependent protein kinase or a part thereof capable of changing cellular localisation upon

activation.

66. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 59, wherein the protein kinase is a calmodulin-dependent serine/threonine protein kinase or a part thereof capable of changing cellular localisation upon activation.
67. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 59, wherein the protein kinase is a mitogen-activated serine/threonine protein kinase or a part thereof capable of changing cellular localisation upon activation.
68. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 67, which codes for an ERK1-F64L-S65T-GFP fusion.
69. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 67, which codes for an EGFP-ERK1 fusion.
70. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 59, wherein the protein kinase is a cyclin-dependent serine/threonine protein kinase or a part thereof capable of changing cellular localisation upon activation.
71. A method according to claim 53 or 54, wherein the nucleic acid construct is a nucleic acid construct according to claim 55 or 56, wherein the biologically active polypeptide is a protein phosphatase or a part thereof capable of changing cellular localisation upon activation.
72. A method according to claim 53 -71, wherein the nucleic acid construct is a nucleic acid construct which is a DNA construct.
73. A method according to claim 53 -72, wherein the nucleic acid construct is a nucleic acid construct according to any of claims 53-72 wherein the gene encoding GFP is derived from *Aequorea victoria*.

74. A method according to claim 73, wherein the nucleic acid construct is a nucleic acid construct according to claim 73 in which the gene encoding GFP is the gene encoding EGFP as defined herein.
75. A method according to claim 73, wherein the nucleic acid construct is a nucleic acid construct according to claim 73 in which the gene encoding a GFP is a gene encoding a GFP variant selected from F64L-GFP, F64L-Y66H-GFP and F64L-S65T-GFP.
76. A method according to claims 72 and 74, wherein the nucleic acid construct is a DNA construct according to claims 72 and 74 or, where applicable, 75, which is a construct as identified by any of the DNA sequences shown in SEQ ID NO: 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, and 152 or is a variant thereof capable of encoding the same fusion polypeptide or a fusion polypeptide which is biologically equivalent thereto, as defined herein.
77. A method comprising a cell containing a nucleic acid construct according to any of claims 53-76 and capable of expressing the sequence encoded by the construct.
78. A method comprising a cell according to claim 77, which is a eukaryotic cell.
79. A method comprising a cell according to claim 77, which is selected from the group consisting of fungal cells, such as yeast cells; invertebrate cells, including insect cells, and vertebrate cells, such as mammalian cells.
80. A method according to any of claims 1-79, as used in a screening program as defined herein.
81. A method according claim 80, wherein the method is a screening program for the identification of a biologically active substance as defined herein that directly or indirectly affects an intracellular signalling pathway and is potentially useful as a medicament, wherein the result of the individual measurement of each substance being screened which indicates its potential biological activity is based on measurement of the redistribution of spatially resolved luminescence in living cells and which undergoes a change in distribution upon activation of an intracellular signalling

pathway.

82. A method according to claim 80, wherein the method is a screening program for the identification of a biologically toxic substance as defined herein that exerts its toxic effect by interfering with an intracellular signalling pathway, wherein the result of the individual measurement of each substance being screened which indicates its potential biologically toxic activity is based on measurement of the redistribution of said fluorescent probe in living cells and which undergoes a change in distribution upon activation of an intracellular signalling pathway.
83. A method according to any of claims 1-82 wherein a fluorescent probe is used in back-tracking of signal transduction pathways as defined herein.
84. A method according to any of claims 1-83, for treating a condition or disease related to the intracellular function of a protein kinase comprising administering to a patient suffering from said condition or disease an effective amount of a compound which has been discovered by any method.
85. A compound that modulates a component of an intracellular pathway as defined herein, as determined by any method according to any of claims 1-83.
86. A medical composition comprising a therapeutic amount of a compound identified according to any method according to any of claims 1-83.
87. A method of selectively treating a patient suffering from an ailment which responds to medical treatment comprising obtaining a primary cells from said patient, transfecting the cells with at least one DNA sequence encoding a fluorescent probe according to any of the preceding claims, culturing the cells under conditions permitting the expression of said probes and exposing it to an array of medicaments suspected of being capable of alleviating said ailment, then comparing changes in fluorescence patterns or redistribution patterns of the fluorescent probes in the intact living cells to detect the cellular response to the specific medicaments (obtaining a cellular action profile), then selecting a medicament(s) based on desired activity and acceptable level of side effects and administering an effective amount of said medicament(s) to said

patient.

88. A method according to any of claims 1-83 of identifying a drug target among the group of biologically active polypeptides that are components of intracellular signalling pathways.

Modtaget PD
15 OKT. 1998

1

Figure 1



a)



b)



c)



d)

Modtaget PD

15 OCT. 1998

Figure 2

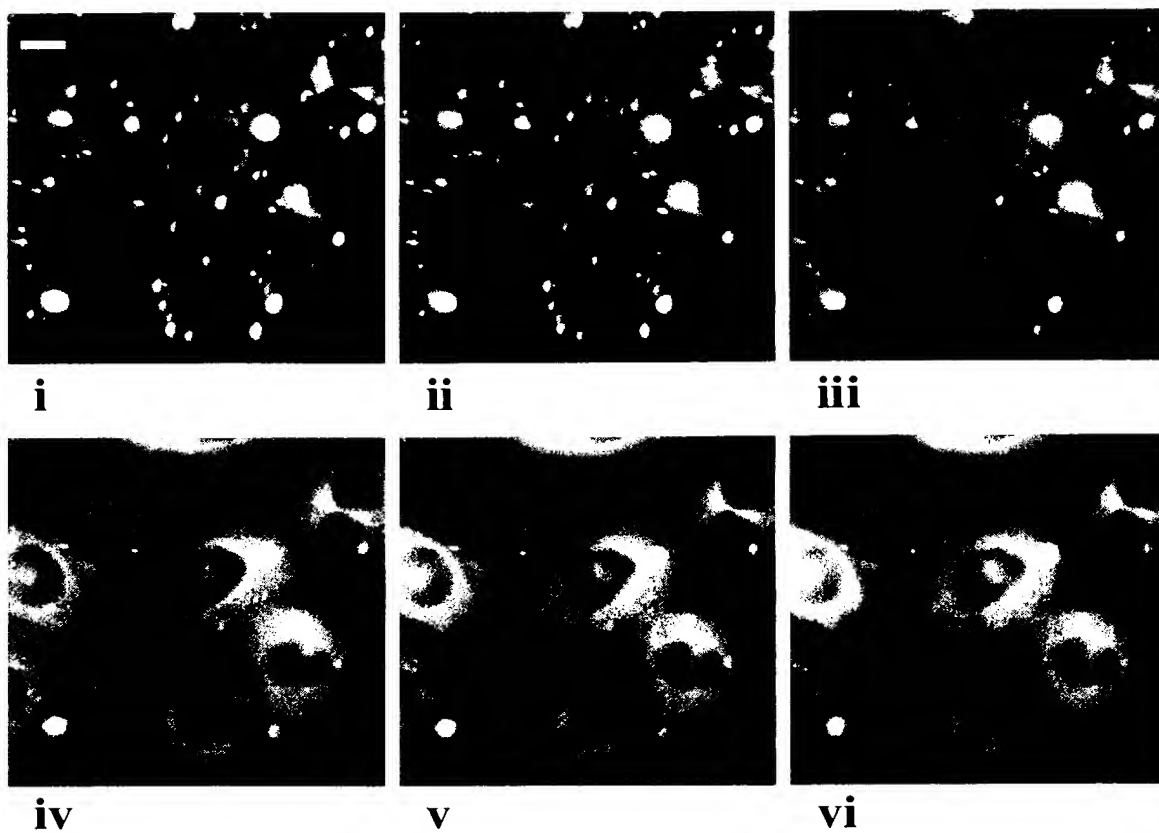
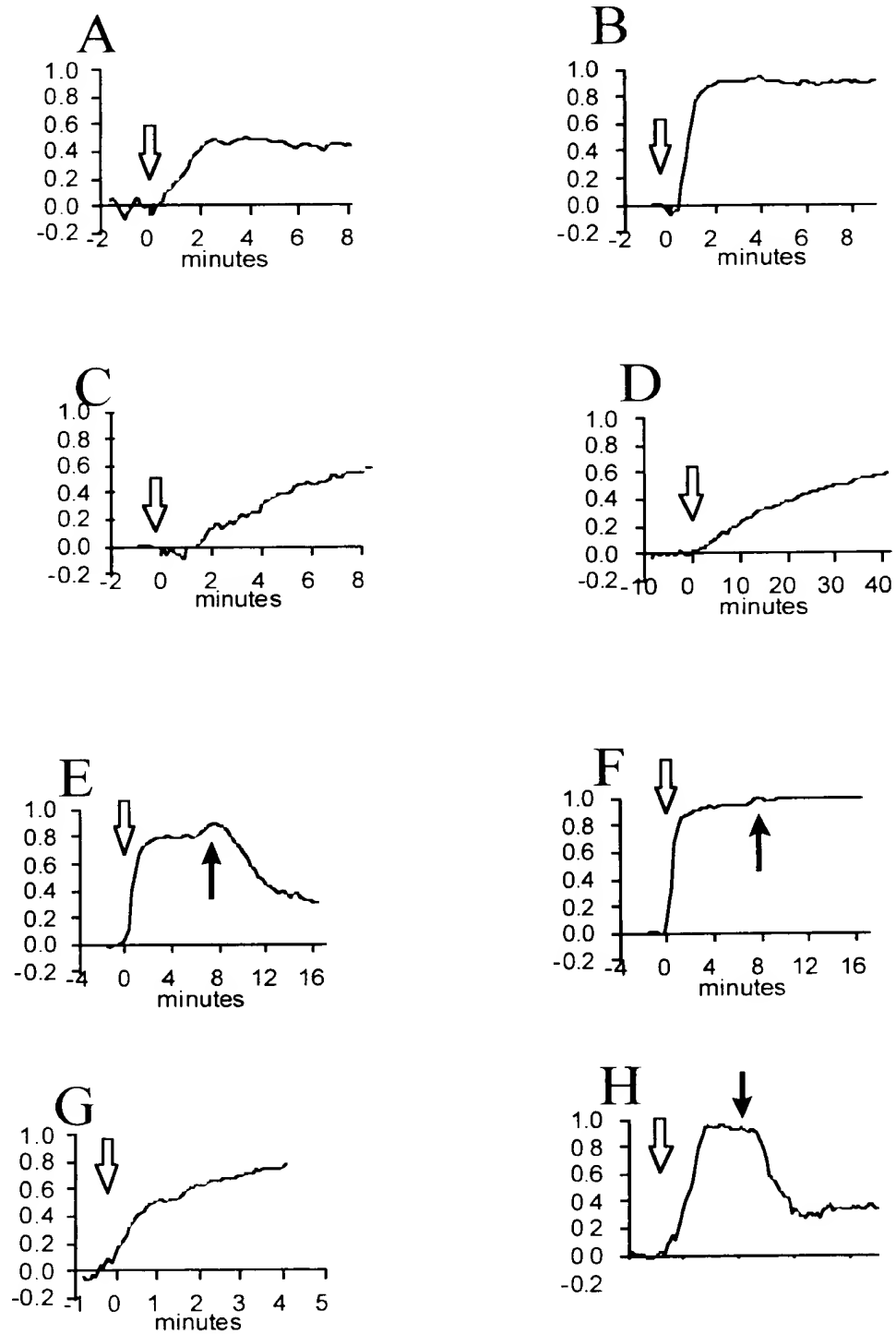
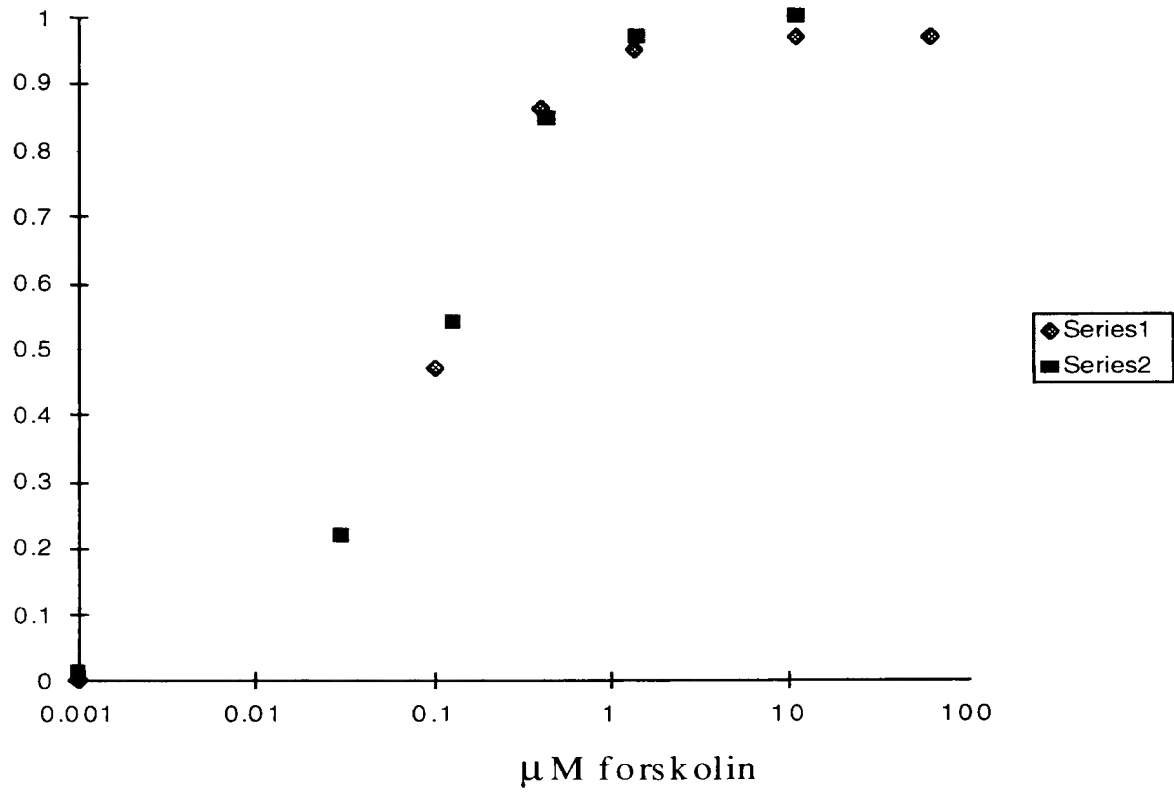


Figure 3



Modtaget PD
15 OKT. 1998

Figure 4

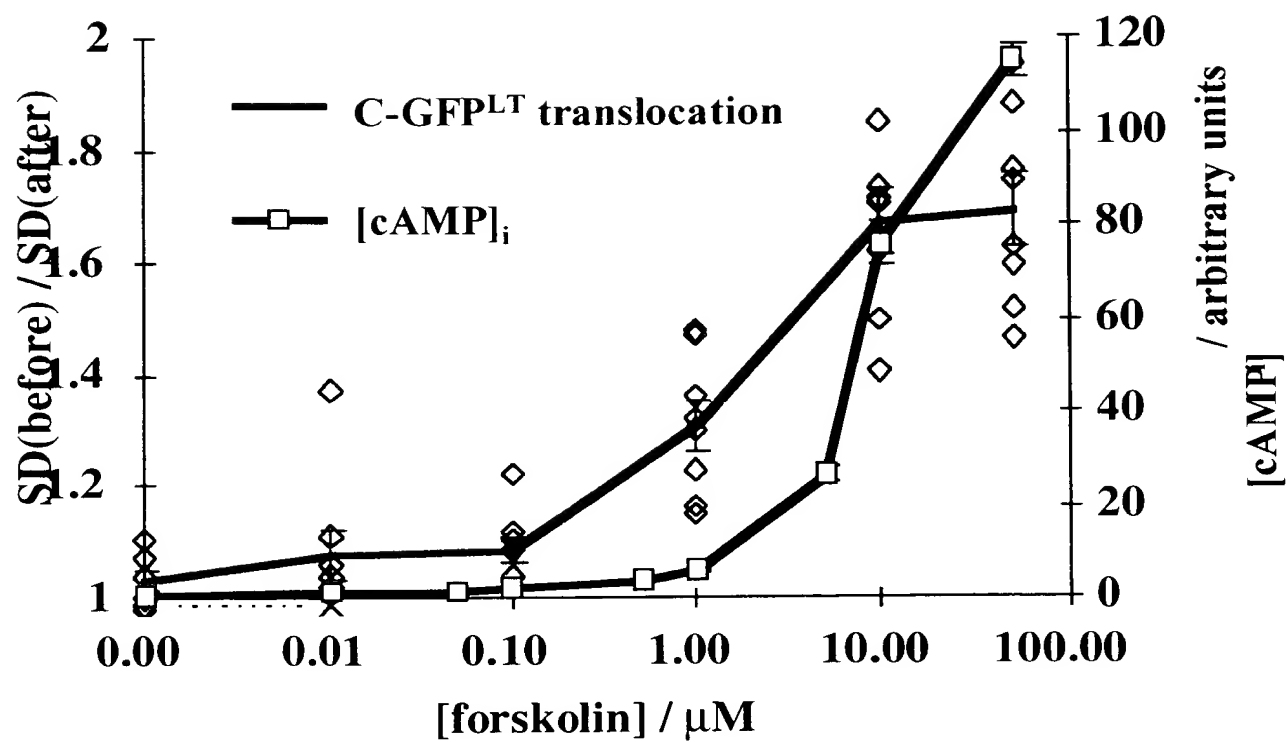


15 OKT. 1998

Figure 5

[forskolin] μ M	$t_{1/2\max}$ / s	t_{\max} / s
1	115 \pm 21	310 \pm 31
10	69 \pm 14	224 \pm 47
50	47 \pm 10	125 \pm 28

Figure 6



Modtaget PD

15 OKT. 1998

7

Figure 7



a)



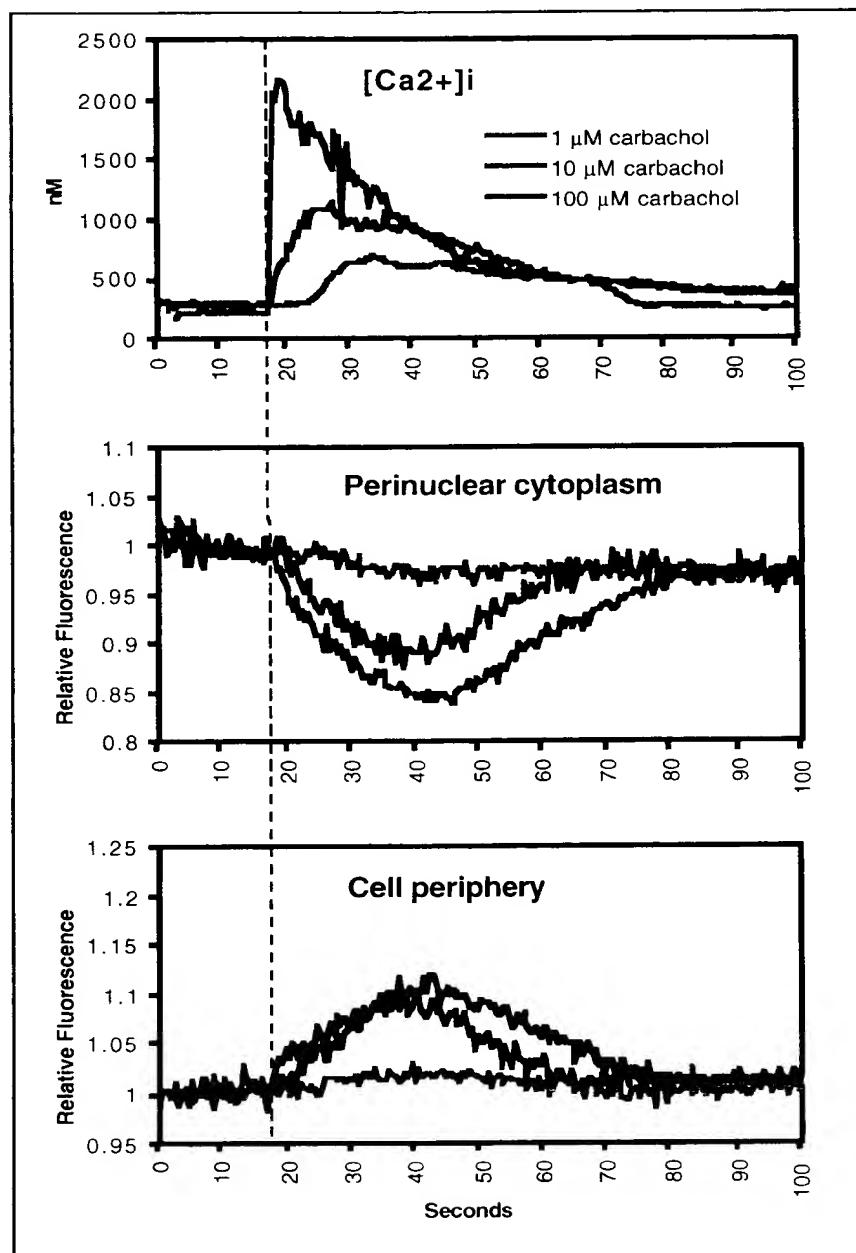
b)



c)

Modtaget PD
15 OKT. 1998

Figure 8



Modtaget PD
15 OKT. 1998

9

Figure 9



Modtaget PD

15 OCT. 1998

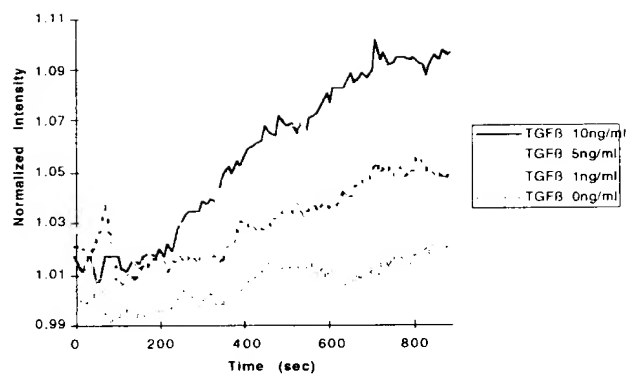
10

Figure 10

a)



b)



c)

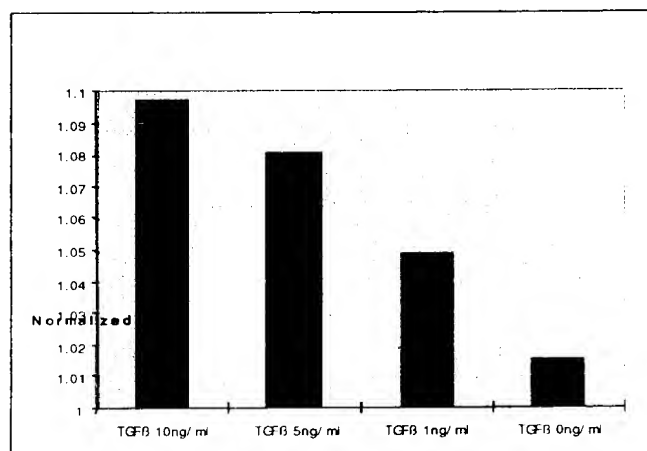
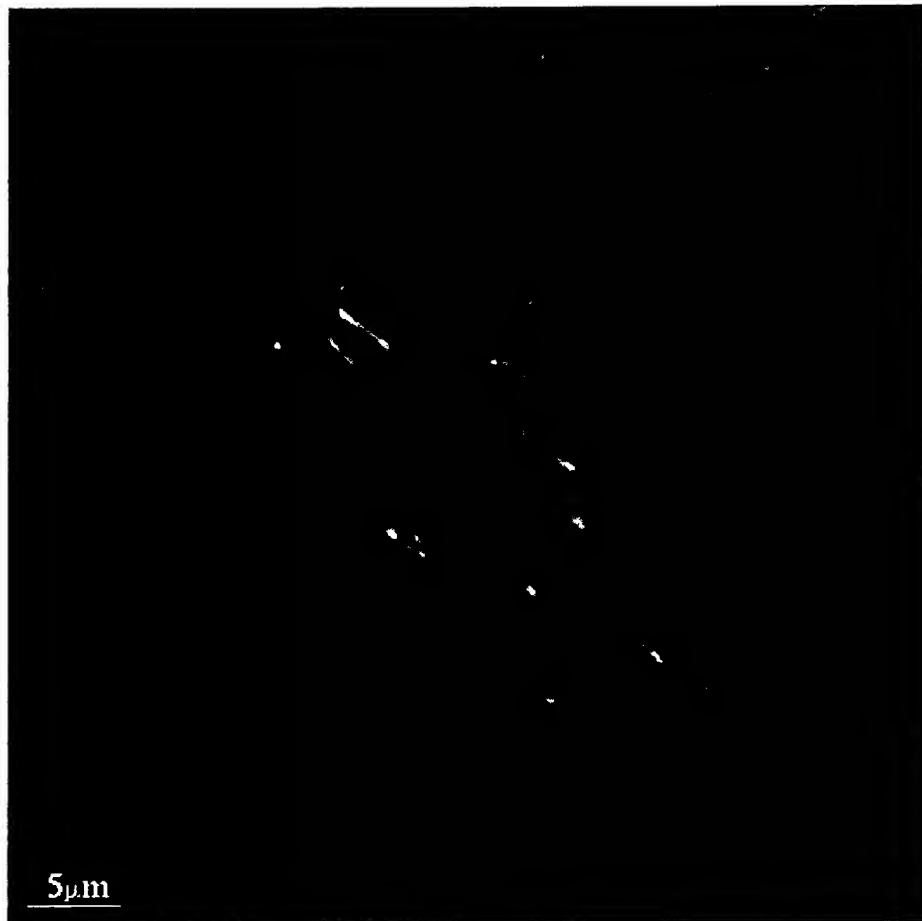


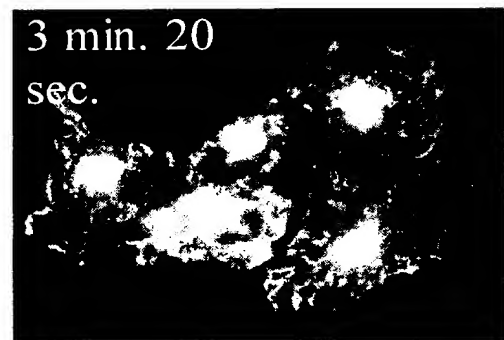
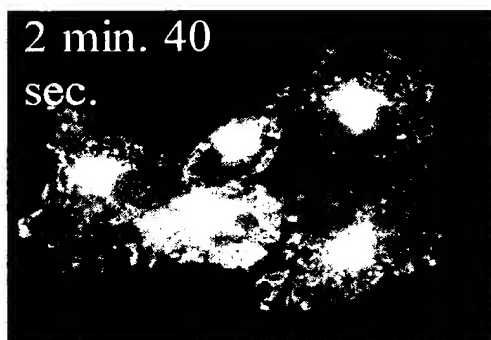
Figure 11



Modtaget PD
15 OKT. 1998

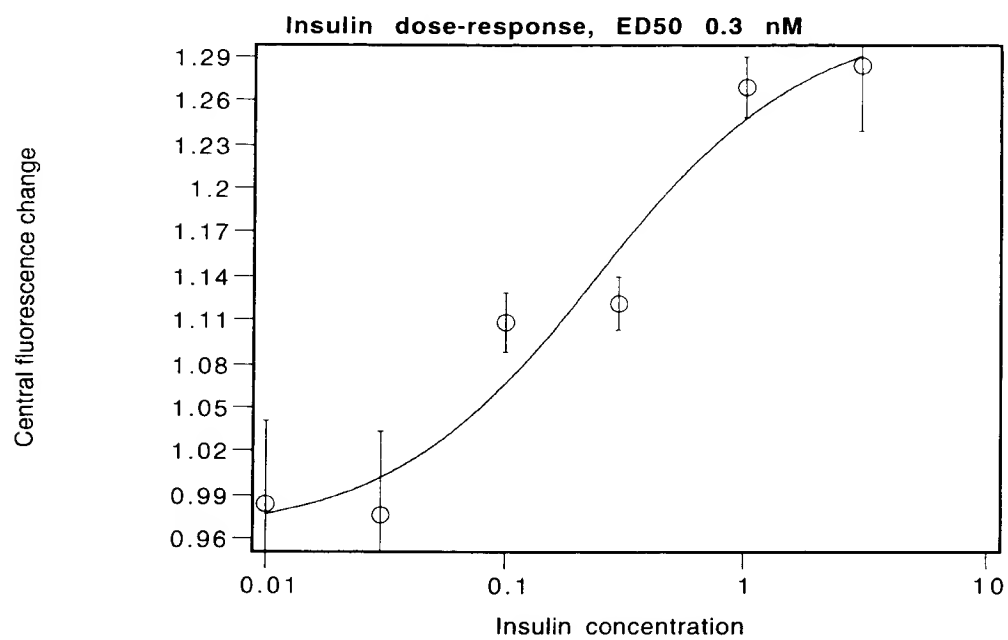
12

Fig. 12



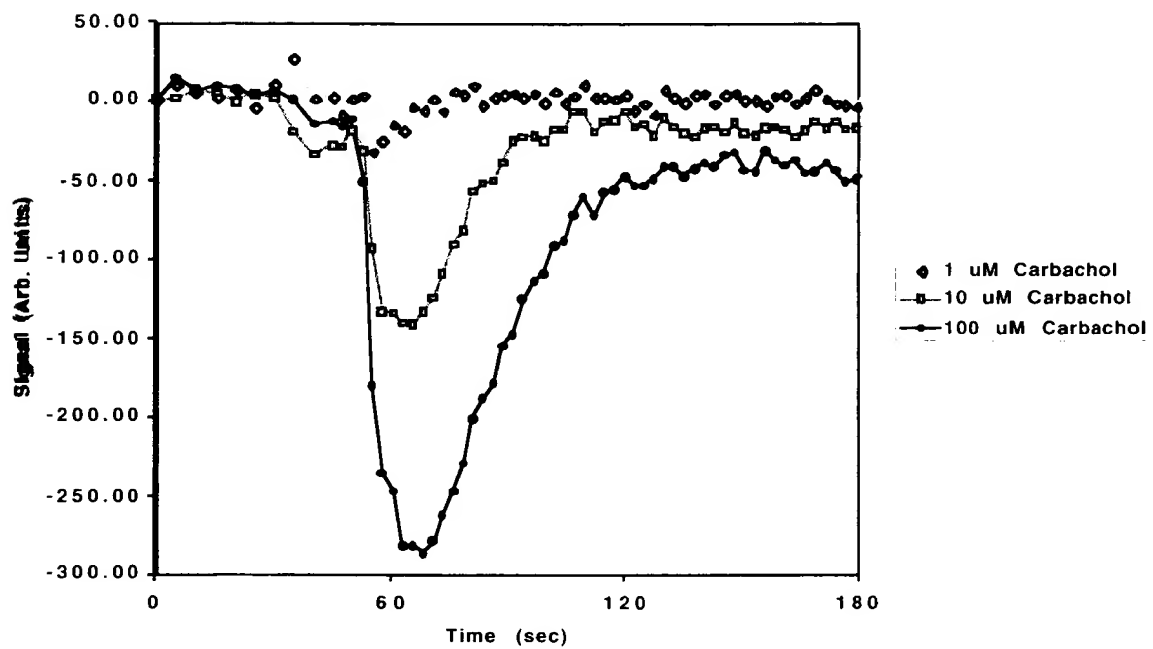
Modtaget PD
15 OKT. 1998

Figure 13



Modtaget PD
15 OKT. 1998

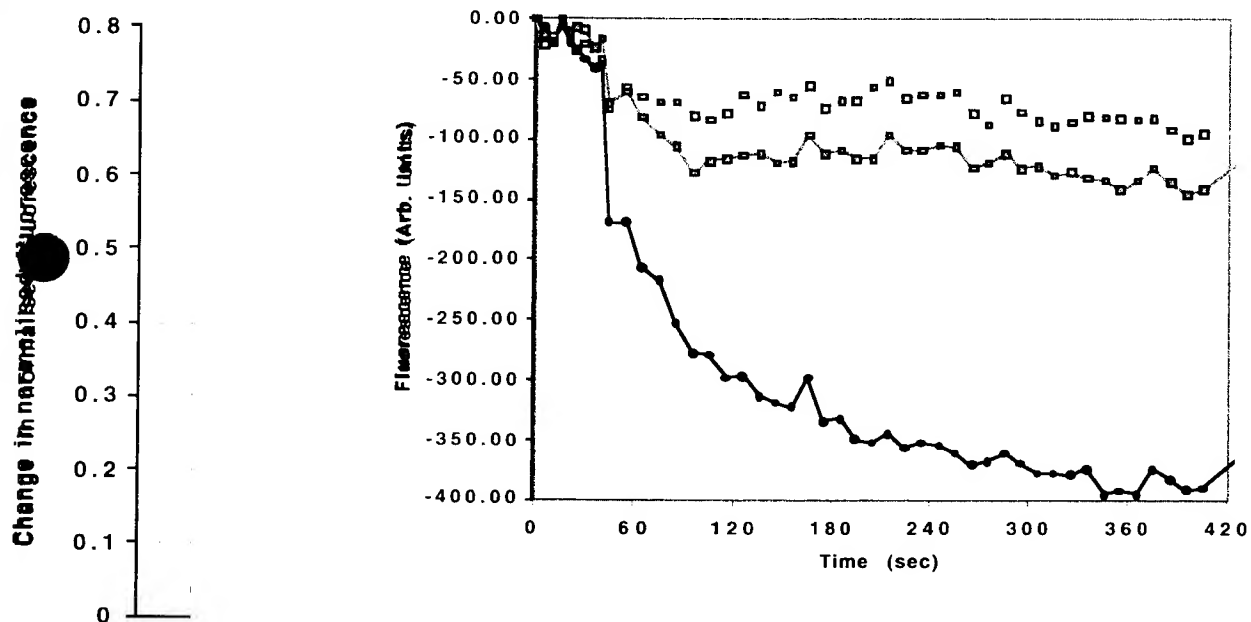
Figure 14



Modtaget F
15 OKT. 1998

Figure 15

Figure 16



Montaget PD
15 OKT. 1998

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: NovoNordisk, BioImage
- (ii) TITLE OF THE INVENTION: An Improved Method of Detecting Cellular Translocation of Biologically Active Polypeptides Using Fluorescence Imaging
- (iii) NUMBER OF SEQUENCES: 165
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: NovoNordisk, BioImage
 - (B) STREET: Mørkhøjbygade 28
 - (C) CITY: Søborg
 - (D) STATE: DK
 - (E) COUNTRY: DENMARK
 - (F) ZIP: 2860
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: , PV&P R
 - (B) REGISTRATION NUMBER:
 - (C) REFERENCE/DOCKET NUMBER:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 53 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTGGACACAA GCTTTGGACA CGGCGCGCCA TGAGTAAAGG AGAAGAACTT TTC

53

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 53 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

2
GTCATCTTCT CGAGTCTTAC TCCTGAGGTT TGTATAGTTC ATCCATGCCA TGT

53

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTGGACACAA GCTTTGGACA CCCTCAGGAT ATGGGCAACG CCGCCGCCGC CAAG

54

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTCATCTTCT CGAGTCTTTC AGGCGCGCCC AAACCTAGTA AACTCCTTGC CACAC

55

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TTGGACACAA GCTTTGGACA CCCTCAGGAT ATGGCTGACG TTTACCCGGC CAACG

55

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTCATCTTCT CGAGTCTTTC AGGCGCGCCC TACTGCACTT TGCAAGATTG GGTGC

55

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 64 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TTGGACACAA GCTTTGGACA CCTCAGGAT ATGGCGGCGG CGGCGGCGGC TCCGGGGGGC 60
GGGG 64

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 55 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTCATCTTCT CGAGTCTTTC AGGCGCGCCC GGGGCCCTCT GGCGCCCTG GCTGG 55

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TAGAATTCAA CCATGGCGGC GGCGCGGCG 30

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TAGGATCCCT AGGGGGCCTC CAGCACTCC 29

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TACTCGAGTA ACCATGGCGG CGGCGGCGGC G 31

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TAGGATCCAT AGATCTGTAT CCTGG 25

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TAGGATCCTT AAGATCTGTA TCCTGG 26

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATCTCGAGGG AAAATGTCTC AGGAGAGG 28

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATGGATCCTC GGACTCCATC TCTTCTTG 28

(2) INFORMATION FOR SEQ ID NO:16:

5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATGGATCCTC AGGACTCCAT CTCTTCTTG

29

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GTCTCGAGCC ATCATGAGCA GAAGCAAG

28

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GTGGATCCCA CTGCTGCACC TGTGCTA

27

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GTGGATCCTC ACTGCTGCAC CTGTGCTA

28

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGCGAATTC GCCACCATGA GTGCTGAGGG GTACCAGTAC 40

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGCGGATCCT GTCGCCTCTG CTGTGCATAT AC 32

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: p85-top-C

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGGAGATCTA TGAGTGCTGA GGGGTACCAG 30

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGGCGGATCC TCATCGCCTC TGCTGTGCAT ATAC 34

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTGAATTCGA CCATGTCGTC CATCTTGCCA TTC

33

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GTGGTACCCA TGACATGCTT GAGCAACGCA C

31

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GTGGTACCTT ATGACATGCT TGAGCAACGC AC

32

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GTGAATTCGT CAATGGAGCT GGAAAACATC G

31

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTGGATCCCT GCTGCTTCCG GTGGAGTTTCG

30

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs

- 8
- (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GTGGATCCCT AGCTGCTTCC GGTGGAGTTC G

31

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

G TAGATCTAC CATGGCGGGC TGGATCCAGG CC

32

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTGGTACCCA TGAGAGGGAG CCTCTGGCAG A

31

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GTGGTACCTC ATGAGAGGGA GCCTCTGGCA G

31

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GTGAATTCAA CCATGGACAA TATGTCTATT ACG

33

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GTGGATCCCA GTCTAAAGGT TGTGGGTCTG C

31

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GTGGATCCTC AGTCTAAAGG TTGTGGGTCT GC

32

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GTCTCGAGGC ACCATGAGCG ACGTGGC

27

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TGGGATCCGA GGCCGTGCTG CTGGCCG

27

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1896 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
 (ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 1...1891
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG	336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
100 105 110	
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC	384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly	
115 120 125	
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC	432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr	
130 135 140	
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC	480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn	
145 150 155 160	
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC	528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser	
165 170 175	

GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190	576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205	624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220	672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser 225 230 235 240	720
GGA CTC AGA TCT CGA GCT CAA GCT TCG AAT TCA ACC ATG GCG GCG GCG Gly Leu Arg Ser Arg Ala Gln Ala Ser Asn Ser Thr Met Ala Ala Ala 245 250 255	768
GCG GCT CAG GGG GGC GGG GGC GGG GAG CCC CGT AGA ACC GAG GGG GTC Ala Ala Gln Gly Gly Gly Gly Gly Glu Pro Arg Arg Thr Glu Gly Val 260 265 270	816
GGC CCG GGG GTC CCG GGG GAG GTG GAG ATG GTG AAG GGG CAG CCG TTC Gly Pro Gly Val Pro Gly Glu Val Glu Met Val Lys Gly Gln Pro Phe 275 280 285	864
GAC GTG GGC CCG CGC TAC ACG CAG TTG CAG TAC ATC GGC GAG GGC GCG Asp Val Gly Pro Arg Tyr Thr Gln Leu Gln Tyr Ile Gly Glu Gly Ala 290 295 300	912
TAC GGC ATG GTC AGC TCG GCC TAT GAC CAC GTG CGC AAG ACT CGC GTG Tyr Gly Met Val Ser Ser Ala Tyr Asp His Val Arg Lys Thr Arg Val 305 310 315 320	960
GCC ATC AAG AAG ATC AGC CCC TTC GAA CAT CAG ACC TAC TGC CAG CGC Ala Ile Lys Lys Ile Ser Pro Phe Glu His Gln Thr Tyr Cys Gln Arg 325 330 335	1008
ACG CTC CGG GAG ATC CAG ATC CTG CTG CGC TTC CGC CAT GAG AAT GTC Thr Leu Arg Glu Ile Gln Ile Leu Leu Arg Phe Arg His Glu Asn Val 340 345 350	1056
ATC GGC ATC CGA GAC ATT CTG CGG GCG TCC ACC CTG GAA GCC ATG AGA Ile Gly Ile Arg Asp Ile Leu Arg Ala Ser Thr Leu Glu Ala Met Arg 355 360 365	1104
GAT GTC TAC ATT GTG CAG GAC CTG ATG GAG ACT GAC CTG TAC AAG TTG Asp Val Tyr Ile Val Gln Asp Leu Met Glu Thr Asp Leu Tyr Lys Leu 370 375 380	1152
CTG AAA AGC CAG CAG CTG AGC AAT GAC CAT ATC TGC TAC TTC CTC TAC Leu Lys Ser Gln Gln Leu Ser Asn Asp His Ile Cys Tyr Phe Leu Tyr 385 390 395 400	1200
CAG ATC CTG CGG GGC CTC AAG TAC ATC CAC TCC GCC AAC GTG CTC CAC Gln Ile Leu Arg Gly Leu Lys Tyr Ile His Ser Ala Asn Val Leu His	1248

405	410	415	
CGA GAT CTA AAG CCC TCC AAC CTG CTC AGC AAC ACC ACC TGC GAC CTT Arg Asp Leu Lys Pro Ser Asn Leu Leu Ser Asn Thr Thr Cys Asp Leu 420 425 430			1296
AAG ATT TGT GAT TTC GGC CTG GCC CGG ATT GCC GAT CCT GAG CAT GAC Lys Ile Cys Asp Phe Gly Leu Ala Arg Ile Ala Asp Pro Glu His Asp 435 440 445			1344
CAC ACC GGC TTC CTG ACG GAG TAT GTG GCT ACG CGC TGG TAC CGG GCC His Thr Gly Phe Leu Thr Glu Tyr Val Ala Thr Arg Trp Tyr Arg Ala 450 455 460			1392
CCA GAG ATC ATG CTG AAC TCC AAG GGC TAT ACC AAG TCC ATC GAC ATC Pro Glu Ile Met Leu Asn Ser Lys Gly Tyr Thr Lys Ser Ile Asp Ile 465 470 475 480			1440
TGG TCT GTG GGC TGC ATT CTG GCT GAG ATG CTC TCT AAC CGG CCC ATC Trp Ser Val Gly Cys Ile Leu Ala Glu Met Leu Ser Asn Arg Pro Ile 485 490 495			1488
TTC CCT GGC AAG CAC TAC CTG GAT CAG CTC AAC CAC ATT CTG GGC ATC Phe Pro Gly Lys His Tyr Leu Asp Gln Leu Asn His Ile Leu Gly Ile 500 505 510			1536
CTG GGC TCC CCA TCC CAG GAG GAC CTG AAT TGT ATC ATC AAC ATG AAG Leu Gly Ser Pro Ser Gln Glu Asp Leu Asn Cys Ile Ile Asn Met Lys 515 520 525			1584
GCC CGA AAC TAC CTA CAG TCT CTG CCC TCC AAG ACC AAG GTG GCT TGG Ala Arg Asn Tyr Leu Gln Ser Leu Pro Ser Lys Thr Lys Val Ala Trp 530 535 540			1632
GCC AAG CTT TTC CCC AAG TCA GAC TCC AAA GCC CTT GAC CTG CTG GAC Ala Lys Leu Phe Pro Lys Ser Asp Ser Lys Ala Leu Asp Leu Leu Asp 545 550 555 560			1680
CGG ATG TTA ACC TTT AAC CCC AAT AAA CGG ATC ACA GTG GAG GAA GCG Arg Met Leu Thr Phe Asn Pro Asn Lys Arg Ile Thr Val Glu Glu Ala 565 570 575			1728
CTG GCT CAC CCC TAC CTG GAG CAG TAC TAT GAC CCG ACG GAT GAG CCA Leu Ala His Pro Tyr Leu Glu Gln Tyr Tyr Asp Pro Thr Asp Glu Pro 580 585 590			1776
GTG GCC GAG GAG CCC TTC ACC TTC GCC ATG GAG CTG GAT GAC CTA CCT Val Ala Glu Glu Pro Phe Thr Phe Ala Met Glu Leu Asp Asp Leu Pro 595 600 605			1824
AAG GAG CGG CTG AAG GAG CTC ATC TTC CAG GAG ACA GCA CGC TTC CAG Lys Glu Arg Leu Lys Glu Leu Ile Phe Gln Glu Thr Ala Arg Phe Gln 610 615 620			1872
CCC GGA GTG CTG GAG GCC C C C C T A G Pro Gly Val Leu Glu Ala Pro 625 630			1896

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 631 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1           5           10           15
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 20           25           30
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 35           40           45
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50           55           60
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 65           70           75           80
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85           90           95
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
100          105          110
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
115          120          125
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
130          135          140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145          150          155          160
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
165          170          175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
180          185          190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
195          200          205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
210          215          220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
225          230          235          240
Gly Leu Arg Ser Arg Ala Gln Ala Ser Asn Ser Thr Met Ala Ala Ala
245          250          255
Ala Ala Gln Gly Gly Gly Gly Gly Glu Pro Arg Arg Thr Glu Gly Val
260          265          270
Gly Pro Gly Val Pro Gly Glu Val Glu Met Val Lys Gly Gln Pro Phe
275          280          285
Asp Val Gly Pro Arg Tyr Thr Gln Leu Gln Tyr Ile Gly Glu Gly Ala
290          295          300
Tyr Gly Met Val Ser Ser Ala Tyr Asp His Val Arg Lys Thr Arg Val
305          310          315          320
Ala Ile Lys Lys Ile Ser Pro Phe Glu His Gln Thr Tyr Cys Gln Arg
325          330          335
Thr Leu Arg Glu Ile Gln Ile Leu Leu Arg Phe Arg His Glu Asn Val
340          345          350

```

Ile Gly Ile Arg Asp Ile Leu Arg Ala Ser Thr Leu Glu Ala Met Arg
 355 360 365
 Asp Val Tyr Ile Val Gln Asp Leu Met Glu Thr Asp Leu Tyr Lys Leu
 370 375 380
 Leu Lys Ser Gln Gln Leu Ser Asn Asp His Ile Cys Tyr Phe Leu Tyr
 385 390 395 400
 Gln Ile Leu Arg Gly Leu Lys Tyr Ile His Ser Ala Asn Val Leu His
 405 410 415
 Arg Asp Leu Lys Pro Ser Asn Leu Leu Ser Asn Thr Thr Cys Asp Leu
 420 425 430
 Lys Ile Cys Asp Phe Gly Leu Ala Arg Ile Ala Asp Pro Glu His Asp
 435 440 445
 His Thr Gly Phe Leu Thr Glu Tyr Val Ala Thr Arg Trp Tyr Arg Ala
 450 455 460
 Pro Glu Ile Met Leu Asn Ser Lys Gly Tyr Thr Lys Ser Ile Asp Ile
 465 470 475 480
 Trp Ser Val Gly Cys Ile Leu Ala Glu Met Leu Ser Asn Arg Pro Ile
 485 490 495
 Phe Pro Gly Lys His Tyr Leu Asp Gln Leu Asn His Ile Leu Gly Ile
 500 505 510
 Leu Gly Ser Pro Ser Gln Glu Asp Leu Asn Cys Ile Ile Asn Met Lys
 515 520 525
 Ala Arg Asn Tyr Leu Gln Ser Leu Pro Ser Lys Thr Lys Val Ala Trp
 530 535 540
 Ala Lys Leu Phe Pro Lys Ser Asp Ser Lys Ala Leu Asp Leu Leu Asp
 545 550 555 560
 Arg Met Leu Thr Phe Asn Pro Asn Lys Arg Ile Thr Val Glu Glu Ala
 565 570 575
 Leu Ala His Pro Tyr Leu Glu Gln Tyr Tyr Asp Pro Thr Asp Glu Pro
 580 585 590
 Val Ala Glu Glu Pro Phe Thr Phe Ala Met Glu Leu Asp Asp Leu Pro
 595 600 605
 Lys Glu Arg Leu Lys Glu Leu Ile Phe Gln Glu Thr Ala Arg Phe Gln
 610 615 620
 Pro Gly Val Leu Glu Ala Pro
 625 630

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1818 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1815
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10 15

GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG	336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
100 105 110	
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC	384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly	
115 120 125	
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC	432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr	
130 135 140	
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC	480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn	
145 150 155 160	
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC	528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser	
165 170 175	
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC	576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly	
180 185 190	
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG	624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu	
195 200 205	
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC	672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe	
210 215 220	
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC	720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser	
225 230 235 240	
GGA CTC AGA TCT CGA GTA ACC ATG GCG GCG GCG GCG GCG GCG GCG CCG	768
Gly Leu Arg Ser Arg Val Thr Met Ala Ala Ala Ala Ala Ala Gly Pro	

245	250	255	
GAG ATG GTC CGC GGG CAG GTG TTC GAC GTG GGG CCG CGC TAC ACT AAT Glu Met Val Arg Gly Gln Val Phe Asp Val Gly Pro Arg Tyr Thr Asn 260 265 270			816
CTC TCG TAC ATC GGA GAA GGC GCC TAC GGC ATG GTT TGT TCT GCT TAT Leu Ser Tyr Ile Gly Glu Gly Ala Tyr Gly Met Val Cys Ser Ala Tyr 275 280 285			864
GAT AAT CTC AAC AAA GTT CGA GTT GCT ATC AAG AAA ATC AGT CCT TTT Asp Asn Leu Asn Lys Val Arg Val Ala Ile Lys Lys Ile Ser Pro Phe 290 295 300			912
GAG CAC CAG ACC TAC TGT CAG AGA ACC CTG AGA GAG ATA AAA ATC CTA Glu His Gln Thr Tyr Cys Gln Arg Thr Leu Arg Glu Ile Lys Ile Leu 305 310 315 320			960
CTG CGC TTC AGA CAT GAG AAC ATC ATC GGC ATC AAT GAC ATC ATC CGG Leu Arg Phe Arg His Glu Asn Ile Ile Gly Ile Asn Asp Ile Ile Arg 325 330 335			1008
GCA CCA ACC ATT GAG CAG ATG AAA GAT GTA TAT ATA GTA CAG GAC CTC Ala Pro Thr Ile Glu Gln Met Lys Asp Val Tyr Ile Val Gln Asp Leu 340 345 350			1056
ATG GAG ACA GAT CTT TAC AAG CTC TTG AAG ACA CAG CAC CTC AGC AAT Met Glu Thr Asp Leu Tyr Lys Leu Leu Lys Thr Gln His Leu Ser Asn 355 360 365			1104
GAT CAT ATC TGC TAT TTT CTT TAT CAG ATC CTG AGA GGA TTA AAG TAT Asp His Ile Cys Tyr Phe Leu Tyr Gln Ile Leu Arg Gly Leu Lys Tyr 370 375 380			1152
ATA CAT TCA GCT AAT GTT CTG CAC CGT GAC CTC AAG CCT TCC AAC CTC Ile His Ser Ala Asn Val Leu His Arg Asp Leu Lys Pro Ser Asn Leu 385 390 395 400			1200
CTG CTG AAC ACC ACT TGT GAT CTC AAG ATC TGT GAC TTT GGC CTT GCC Leu Leu Asn Thr Thr Cys Asp Leu Lys Ile Cys Asp Phe Gly Leu Ala 405 410 415			1248
CGT GTT GCA GAT CCA GAC CAT GAT CAT ACA GGG TTC TTG ACA GAG TAT Arg Val Ala Asp Pro Asp His Asp His Thr Gly Phe Leu Thr Glu Tyr 420 425 430			1296
GTA GCC ACG CGT TGG TAC AGA GCT CCA GAA ATT ATG TTG AAT TCC AAG Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu Ile Met Leu Asn Ser Lys 435 440 445			1344
GGT TAT ACC AAG TCC ATT GAT ATT TGG TCT GTG GGC TGC ATC CTG GCA Gly Tyr Thr Lys Ser Ile Asp Ile Trp Ser Val Gly Cys Ile Leu Ala 450 455 460			1392
GAG ATG CTA TCC AAC AGG CCT ATC TTC CCA GGA AAG CAT TAC CTT GAC Glu Met Leu Ser Asn Arg Pro Ile Phe Pro Gly Lys His Tyr Leu Asp 465 470 475 480			1440

CAG CTG AAT CAC ATC CTG GGT ATT CTT GGA TCT CCA TCA CAG GAA GAT	1488
Gln Leu Asn His Ile Leu Gly Ile Leu Gly Ser Pro Ser Gln Glu Asp	
485 490 495	
CTG AAT TGT ATA ATA AAT TTA AAA GCT AGA AAC TAT TTG CTT TCT CTC	1536
Leu Asn Cys Ile Ile Asn Leu Lys Ala Arg Asn Tyr Leu Leu Ser Leu	
500 505 510	
CCG CAC AAA AAT AAG GTG CCG TGG AAC AGG TTG TTC CCA AAC GCT GAC	1584
Pro His Lys Asn Lys Val Pro Trp Asn Arg Leu Phe Pro Asn Ala Asp	
515 520 525	
TCC AAA GCT CTG GAT TTA CTG GAT AAA ATG TTG ACA TTT AAC CCT CAC	1632
Ser Lys Ala Leu Asp Leu Leu Asp Lys Met Leu Thr Phe Asn Pro His	
530 535 540	
AAG AGG ATT GAA GTT GAA CAG GCT CTG GCC CAC CCG TAC CTG GAG CAG	1680
Lys Arg Ile Glu Val Glu Gln Ala Leu Ala His Pro Tyr Leu Glu Gln	
545 550 555 560	
TAT TAT GAC CCA AGT GAT GAG CCC ATT GCT GAA GCA CCA TTC AAG TTT	1728
Tyr Tyr Asp Pro Ser Asp Glu Pro Ile Ala Glu Ala Pro Phe Lys Phe	
565 570 575	
GAC ATG GAG CTG GAC GAC TTA CCT AAG GAG AAG CTC AAA GAA CTC ATT	1776
Asp Met Glu Leu Asp Asp Leu Pro Lys Glu Lys Leu Lys Glu Leu Ile	
580 585 590	
TTT GAA GAG ACT GCT CGA TTC CAG CCA GGA TAC AGA TCT TAA	1818
Phe Glu Glu Thr Ala Arg Phe Gln Pro Gly Tyr Arg Ser	
595 600 605	

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 605 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	

Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	100	105	110
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	115	120	125
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	130	135	140
Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	145	150	155
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	165	170	175
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	180	185	190
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	195	200	205
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	210	215	220
Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Ser	225	230	235
Gly	Leu	Arg	Ser	Arg	Val	Thr	Met	Ala	Ala	Ala	Ala	Ala	Ala	Gly	Pro	245	250	255
Glu	Met	Val	Arg	Gly	Gln	Val	Phe	Asp	Val	Gly	Pro	Arg	Tyr	Thr	Asn	260	265	270
Leu	Ser	Tyr	Ile	Gly	Glu	Gly	Ala	Tyr	Gly	Met	Val	Cys	Ser	Ala	Tyr	275	280	285
Asp	Asn	Leu	Asn	Lys	Val	Arg	Val	Ala	Ile	Lys	Lys	Ile	Ser	Pro	Phe	290	295	300
Glu	His	Gln	Thr	Tyr	Cys	Gln	Arg	Thr	Leu	Arg	Glu	Ile	Lys	Ile	Leu	305	310	315
Leu	Arg	Phe	Arg	His	Glu	Asn	Ile	Ile	Gly	Ile	Asn	Asp	Ile	Ile	Arg	325	330	335
Ala	Pro	Thr	Ile	Glu	Gln	Met	Lys	Asp	Val	Tyr	Ile	Val	Gln	Asp	Leu	340	345	350
Met	Glu	Thr	Asp	Leu	Tyr	Lys	Leu	Leu	Lys	Thr	Gln	His	Leu	Ser	Asn	355	360	365
Asp	His	Ile	Cys	Tyr	Phe	Leu	Tyr	Gln	Ile	Leu	Arg	Gly	Leu	Lys	Tyr	370	375	380
Ile	His	Ser	Ala	Asn	Val	Leu	His	Arg	Asp	Leu	Lys	Pro	Ser	Asn	Leu	385	390	395
Leu	Leu	Asn	Thr	Thr	Cys	Asp	Leu	Lys	Ile	Cys	Asp	Phe	Gly	Leu	Ala	405	410	415
Arg	Val	Ala	Asp	Pro	Asp	His	Asp	His	Thr	Gly	Phe	Leu	Thr	Glu	Tyr	420	425	430
Val	Ala	Thr	Arg	Trp	Tyr	Arg	Ala	Pro	Glu	Ile	Met	Leu	Asn	Ser	Lys	435	440	445
Gly	Tyr	Thr	Lys	Ser	Ile	Asp	Ile	Trp	Ser	Val	Gly	Cys	Ile	Leu	Ala	450	455	460
Glu	Met	Leu	Ser	Asn	Arg	Pro	Ile	Phe	Pro	Gly	Lys	His	Tyr	Leu	Asp	465	470	475
Gln	Leu	Asn	His	Ile	Leu	Gly	Ile	Leu	Gly	Ser	Pro	Ser	Gln	Glu	Asp	485	490	495
Leu	Asn	Cys	Ile	Ile	Asn	Leu	Lys	Ala	Arg	Asn	Tyr	Leu	Leu	Ser	Leu	500	505	510
Pro	His	Lys	Asn	Lys	Val	Pro	Trp	Asn	Arg	Leu	Phe	Pro	Asn	Ala	Asp	515	520	525
Ser	Lys	Ala	Leu	Asp	Leu	Leu	Asp	Lys	Met	Leu	Thr	Phe	Asn	Pro	His	530	535	540
Lys	Arg	Ile	Glu	Val	Glu	Gln	Ala	Leu	Ala	His	Pro	Tyr	Leu	Glu	Gln	545	550	555
																		560

Tyr Tyr Asp Pro Ser Asp Glu Pro Ile Ala Glu Ala Pro Phe Lys Phe
565 570 575
Asp Met Glu Leu Asp Asp Leu Pro Lys Glu Lys Leu Lys Glu Leu Ile
580 585 590
Phe Glu Glu Thr Ala Arg Phe Gln Pro Gly Tyr Arg Ser
595 600 605

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2529 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2526
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG	336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
100 105 110	
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC	384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly	
115 120 125	
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC	432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr	

130	135	140	
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn 145 150 155 160			480
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser 165 170 175			528
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190			576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205			624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220			672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser 225 230 235 240			720
GGA CTC AGA TCT CGA GCT CAA GCT TCG AAT TCG TCA ATG GAG CTG GAA Gly Leu Arg Ser Arg Ala Gln Ala Ser Asn Ser Ser Met Glu Leu Glu 245 250 255			768
AAC ATC GTG GCC AAC ACG GTC TTG CTG AAA GCC AGG GAA GGG GGC GGA Asn Ile Val Ala Asn Thr Val Leu Leu Lys Ala Arg Glu Gly Gly Gly 260 265 270			816
GGA AAG CGC AAA GGG AAA AGC AAG AAG TGG AAA GAA ATC CTG AAG TTC Gly Lys Arg Lys Gly Lys Ser Lys Lys Trp Lys Glu Ile Leu Lys Phe 275 280 285			864
CCT CAC ATT AGC CAG TGT GAA GAC CTC CGA AGG ACC ATA GAC AGA GAT Pro His Ile Ser Gln Cys Glu Asp Leu Arg Arg Thr Ile Asp Arg Asp 290 295 300			912
TAC TGC AGT TTA TGT GAC AAG CAG CCA ATC GGG AGG CTG CTT TTC CGG Tyr Cys Ser Leu Cys Asp Lys Gln Pro Ile Gly Arg Leu Leu Phe Arg 305 310 315 320			960
CAG TTT TGT GAA ACC AGG CCT GGG CTG GAG TGT TAC ATT CAG TTC CTG Gln Phe Cys Glu Thr Arg Pro Gly Leu Glu Cys Tyr Ile Gln Phe Leu 325 330 335			1008
GAC TCC GTG GCA GAA TAT GAA GTT ACT CCA GAT GAA AAA CTG GGA GAG Asp Ser Val Ala Glu Tyr Glu Val Thr Pro Asp Glu Lys Leu Gly Glu 340 345 350			1056
AAA GGG AAG GAA ATT ATG ACC AAG TAC CTC ACC CCA AAG TCC CCT GTT Lys Gly Lys Glu Ile Met Thr Lys Tyr Leu Thr Pro Lys Ser Pro Val 355 360 365			1104

TTC ATA GCC CAA GTT GGC CAA GAC CTG GTC TCC CAG ACG GAG GAG AAG	1152
Phe Ile Ala Gln Val Gly Gln Asp Leu Val Ser Gln Thr Glu Glu Lys	
370 375 380	
CTC CTA CAG AAG CCG TGC AAA GAA CTC TTT TCT GCC TGT GCA CAG TCT	1200
Leu Leu Gln Lys Pro Cys Lys Glu Leu Phe Ser Ala Cys Ala Gln Ser	
385 390 395 400	
GTC CAC GAG TAC CTG AGG GGA GAA CCA TTC CAC GAA TAT CTG GAC AGC	1248
Val His Glu Tyr Leu Arg Gly Glu Pro Phe His Glu Tyr Leu Asp Ser	
405 410 415	
ATG TTT TTT GAC CGC TTT CTC CAG TGG AAG TGG TTG GAA AGG CAA CCG	1296
Met Phe Phe Asp Arg Phe Leu Gln Trp Lys Trp Leu Glu Arg Gln Pro	
420 425 430	
GTG ACC AAA AAC ACT TTC AGG CAG TAT CGA GTG CTA GGA AAA GGG GGC	1344
Val Thr Lys Asn Thr Phe Arg Gln Tyr Arg Val Leu Gly Lys Gly Gly	
435 440 445	
TTC GGG GAG GTC TGT GCC TGC CAG GTT CGG GCC ACG GGT AAA ATG TAT	1392
Phe Gly Glu Val Cys Ala Cys Gln Val Arg Ala Thr Gly Lys Met Tyr	
450 455 460	
GCC TGC AAG CGC TTG GAG AAG AAG AGG ATC AAA AAG AGG AAA GGG GAG	1440
Ala Cys Lys Arg Leu Glu Lys Lys Arg Ile Lys Lys Arg Lys Gly Glu	
465 470 475 480	
TCC ATG GCC CTC AAT GAG AAG CAG ATC CTC GAG AAG GTC AAC AGT CAG	1488
Ser Met Ala Leu Asn Glu Lys Gln Ile Leu Glu Lys Val Asn Ser Gln	
485 490 495	
TTT GTG GTC AAC CTG GCC TAT GCC TAC GAG ACC AAG GAT GCA CTG TGC	1536
Phe Val Val Asn Leu Ala Tyr Ala Tyr Glu Thr Lys Asp Ala Leu Cys	
500 505 510	
TTG GTC CTG ACC ATC ATG AAT GGG GGT GAC CTG AAG TTC CAC ATC TAC	1584
Leu Val Leu Thr Ile Met Asn Gly Gly Asp Leu Lys Phe His Ile Tyr	
515 520 525	
AAC ATG GGC AAC CCT GGC TTC GAG GAG GAG CGG GCC TTG TTT TAT GCG	1632
Asn Met Gly Asn Pro Gly Phe Glu Glu Glu Arg Ala Leu Phe Tyr Ala	
530 535 540	
GCA GAG ATC CTC TGC GGC TTA GAA GAC CTC CAC CGT GAG AAC ACC GTC	1680
Ala Glu Ile Leu Cys Gly Leu Glu Asp Leu His Arg Glu Asn Thr Val	
545 550 555 560	
TAC CGA GAT CTG AAA CCT GAA AAC ATC CTG TTA GAT GAT TAT GGC CAC	1728
Tyr Arg Asp Leu Lys Pro Glu Asn Ile Leu Leu Asp Asp Tyr Gly His	
565 570 575	
ATT AGG ATC TCA GAC CTG GGC TTG GCT GTG AAG ATC CCC GAG GGA GAC	1776
Ile Arg Ile Ser Asp Leu Gly Leu Ala Val Lys Ile Pro Glu Gly Asp	
580 585 590	
CTG ATC CGC GGC CGG GTG GGC ACT GTT GGC TAC ATG GCC CCC GAA GTC	1824
Leu Ile Arg Gly Arg Val Gly Thr Val Gly Tyr Met Ala Pro Glu Val	

595	600	605	
CTG AAC AAC CAG AGG TAC GGC CTG AGC CCC GAC TAC TGG GGC CTT GGC Leu Asn Asn Gln Arg Tyr Gly Leu Ser Pro Asp Tyr Trp Gly Leu Gly 610 615 620			1872
TGC CTC ATC TAT GAG ATG ATC GAG GGC CAG TCG CCG TTC CGC GGC CGT Cys Leu Ile Tyr Glu Met Ile Glu Gly Gln Ser Pro Phe Arg Gly Arg 625 630 635 640			1920
AAG GAG AAG GTG AAG CGG GAG GAG GTG GAC CGC CGG GTC CTG GAG ACG Lys Glu Lys Val Lys Arg Glu Glu Val Asp Arg Arg Val Leu Glu Thr 645 650 655			1968
GAG GAG GTG TAC TCC CAC AAG TTC TCC GAG GAG GCC AAG TCC ATC TGC Glu Glu Val Tyr Ser His Lys Phe Ser Glu Glu Ala Lys Ser Ile Cys 660 665 670			2016
AAG ATG CTG CTC ACG AAA GAT GCG AAG CAG AGG CTG GGC TGC CAG GAG Lys Met Leu Leu Thr Lys Asp Ala Lys Gln Arg Leu Gly Cys Gln Glu 675 680 685			2064
GAG GGG GCT GCA GAG GTC AAG AGA CAC CCC TTC TTC AGG AAC ATG AAC Glu Gly Ala Ala Glu Val Lys Arg His Pro Phe Phe Arg Asn Met Asn 690 695 700			2112
TTC AAG CGC TTA GAA GCC GGG ATG TTG GAC CCT CCC TTC GTT CCA GAC Phe Lys Arg Leu Glu Ala Gly Met Leu Asp Pro Pro Phe Val Pro Asp 705 710 715 720			2160
CCC CGC GCT GTG TAC TGT AAG GAC GTG CTG GAC ATC GAG CAG TTC TCC Pro Arg Ala Val Tyr Cys Lys Asp Val Leu Asp Ile Glu Gln Phe Ser 725 730 735			2208
ACT GTG AAG GGC GTC AAT CTG GAC CAC ACA GAC GAC GAC TTC TAC TCC Thr Val Lys Gly Val Asn Leu Asp His Thr Asp Asp Asp Phe Tyr Ser 740 745 750			2256
AAG TTC TCC ACG GGC TCT GTG TCC ATC CCA TGG CAA AAC GAG ATG ATA Lys Phe Ser Thr Gly Ser Val Ser Ile Pro Trp Gln Asn Glu Met Ile 755 760 765			2304
GAA ACA GAA TGC TTT AAG GAG CTG AAC GTG TTT GGA CCT AAT GGT ACC Glu Thr Glu Cys Phe Lys Glu Leu Asn Val Phe Gly Pro Asn Gly Thr 770 775 780			2352
CTC CCG CCA GAT CTG AAC AGA AAC CAC CCT CCG GAA CCG CCC AAG AAA Leu Pro Pro Asp Leu Asn Arg Asn His Pro Pro Glu Pro Pro Lys Lys 785 790 795 800			2400
GGG CTG CTC CAG AGA CTC TTC AAG CGG CAG CAT CAG AAC AAT TCC AAG Gly Leu Leu Gln Arg Leu Phe Lys Arg Gln His Gln Asn Asn Ser Lys 805 810 815			2448
AGT TCG CCC AGC TCC AAG ACC AGT TTT AAC CAC CAC ATA AAC TCA AAC Ser Ser Pro Ser Ser Lys Thr Ser Phe Asn His His Ile Asn Ser Asn 820 825 830			2496

CAT GTC AGC TCG AAC TCC ACC GGA AGC AGC TAG
 His Val Ser Ser Asn Ser Thr Gly Ser Ser
 835 840

2529

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 842 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	1	5	10	15
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	20	25	30	
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	35	40	45	
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	50	55	60	
Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	65	70	75	80
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	85	90	95	
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	100	105	110	
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	115	120	125	
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	130	135	140	
Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	145	150	155	160
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	165	170	175	
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	180	185	190	
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	195	200	205	
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	210	215	220	
Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Ser	225	230	235	240
Gly	Leu	Arg	Ser	Arg	Ala	Gln	Ala	Ser	Asn	Ser	Ser	Met	Glu	Leu	Glu	245	250	255	
Asn	Ile	Val	Ala	Asn	Thr	Val	Leu	Leu	Lys	Ala	Arg	Glu	Gly	Gly	Gly	260	265	270	
Gly	Lys	Arg	Lys	Gly	Lys	Ser	Lys	Lys	Trp	Lys	Glu	Ile	Leu	Lys	Phe	275	280	285	
Pro	His	Ile	Ser	Gln	Cys	Glu	Asp	Leu	Arg	Arg	Thr	Ile	Asp	Arg	Asp	290	295	300	
Tyr	Cys	Ser	Leu	Cys	Asp	Lys	Gln	Pro	Ile	Gly	Arg	Leu	Leu	Phe	Arg	305	310	315	320

Gln Phe Cys Glu Thr Arg Pro Gly Leu Glu Cys Tyr Ile Gln Phe Leu
 325 330 335
 Asp Ser Val Ala Glu Tyr Glu Val Thr Pro Asp Glu Lys Leu Gly Glu
 340 345 350
 Lys Gly Lys Glu Ile Met Thr Lys Tyr Leu Thr Pro Lys Ser Pro Val
 355 360 365
 Phe Ile Ala Gln Val Gly Gln Asp Leu Val Ser Gln Thr Glu Glu Lys
 370 375 380
 Leu Leu Gln Lys Pro Cys Lys Glu Leu Phe Ser Ala Cys Ala Gln Ser
 385 390 395 400
 Val His Glu Tyr Leu Arg Gly Glu Pro Phe His Glu Tyr Leu Asp Ser
 405 410 415
 Met Phe Phe Asp Arg Phe Leu Gln Trp Lys Trp Leu Glu Arg Gln Pro
 420 425 430
 Val Thr Lys Asn Thr Phe Arg Gln Tyr Arg Val Leu Gly Lys Gly Gly
 435 440 445
 Phe Gly Glu Val Cys Ala Cys Gln Val Arg Ala Thr Gly Lys Met Tyr
 450 455 460
 Ala Cys Lys Arg Leu Glu Lys Lys Arg Ile Lys Lys Arg Lys Gly Glu
 465 470 475 480
 Ser Met Ala Leu Asn Glu Lys Gln Ile Leu Glu Lys Val Asn Ser Gln
 485 490 495
 Phe Val Val Asn Leu Ala Tyr Ala Tyr Glu Thr Lys Asp Ala Leu Cys
 500 505 510
 Leu Val Leu Thr Ile Met Asn Gly Gly Asp Leu Lys Phe His Ile Tyr
 515 520 525
 Asn Met Gly Asn Pro Gly Phe Glu Glu Glu Arg Ala Leu Phe Tyr Ala
 530 535 540
 Ala Glu Ile Leu Cys Gly Leu Glu Asp Leu His Arg Glu Asn Thr Val
 545 550 555 560
 Tyr Arg Asp Leu Lys Pro Glu Asn Ile Leu Leu Asp Asp Tyr Gly His
 565 570 575
 Ile Arg Ile Ser Asp Leu Gly Leu Ala Val Lys Ile Pro Glu Gly Asp
 580 585 590
 Leu Ile Arg Gly Arg Val Gly Thr Val Gly Tyr Met Ala Pro Glu Val
 595 600 605
 Leu Asn Asn Gln Arg Tyr Gly Leu Ser Pro Asp Tyr Trp Gly Leu Gly
 610 615 620
 Cys Leu Ile Tyr Glu Met Ile Glu Gly Gln Ser Pro Phe Arg Gly Arg
 625 630 635 640
 Lys Glu Lys Val Lys Arg Glu Glu Val Asp Arg Arg Val Leu Glu Thr
 645 650 655
 Glu Glu Val Tyr Ser His Lys Phe Ser Glu Glu Ala Lys Ser Ile Cys
 660 665 670
 Lys Met Leu Leu Thr Lys Asp Ala Lys Gln Arg Leu Gly Cys Gln Glu
 675 680 685
 Glu Gly Ala Ala Glu Val Lys Arg His Pro Phe Phe Arg Asn Met Asn
 690 695 700
 Phe Lys Arg Leu Glu Ala Gly Met Leu Asp Pro Pro Phe Val Pro Asp
 705 710 715 720
 Pro Arg Ala Val Tyr Cys Lys Asp Val Leu Asp Ile Glu Gln Phe Ser
 725 730 735
 Thr Val Lys Gly Val Asn Leu Asp His Thr Asp Asp Asp Phe Tyr Ser
 740 745 750
 Lys Phe Ser Thr Gly Ser Val Ser Ile Pro Trp Gln Asn Glu Met Ile
 755 760 765
 Glu Thr Glu Cys Phe Lys Glu Leu Asn Val Phe Gly Pro Asn Gly Thr
 770 775 780

Leu	Pro	Pro	Asp	Leu	Asn	Arg	Asn	His	Pro	Pro	Glu	Pro	Pro	Lys	Lys
785					790					795					800
Gly	Leu	Leu	Gln	Arg	Leu	Phe	Lys	Arg	Gln	His	Gln	Asn	Asn	Ser	Lys
				805					810					815	
Ser	Ser	Pro	Ser	Ser	Lys	Thr	Ser	Phe	Asn	His	His	Ile	Asn	Ser	Asn
			820					825					830		
His	Val	Ser	Ser	Asn	Ser	Thr	Gly	Ser	Ser						
		835					840								

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1902 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...1899
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	48
Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
1				5				10						15		
GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	96
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			20					25					30			
GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	AAG	TTC	ATC	144
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
		35					40					45				
TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	192
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
	50					55					60					
CTG	ACC	TAC	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	CAC	ATG	AAG	240
Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	
65				70					75						80	
CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	288
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	
				85					90					95		
CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	336
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	
			100					105					110			
GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	CTG	AAG	GGC	384
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	
		115					120					125				

ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr 130 135 140	432
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn 145 150 155 160	480
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser 165 170 175	528
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190	576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205	624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220	672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser 225 230 235 240	720
GGA CTC AGA TCT CGA GCT CGA GCC ATC ATG AGC AGA AGC AAG CGT GAC Gly Leu Arg Ser Arg Ala Arg Ala Ile Met Ser Arg Ser Lys Arg Asp 245 250 255	768
AAC AAT TTT TAT AGT GTA GAG ATT GGA GAT TCT ACA TTC ACA GTC CTG Asn Asn Phe Tyr Ser Val Glu Ile Gly Asp Ser Thr Phe Thr Val Leu 260 265 270	816
AAA CGA TAT CAG AAT TTA AAA CCT ATA GGC TCA GGA GCT CAA GGA ATA Lys Arg Tyr Gln Asn Leu Lys Pro Ile Gly Ser Gly Ala Gln Gly Ile 275 280 285	864
GTA TGC GCA GCT TAT GAT GCC ATT CTT GAA AGA AAT GTT GCA ATC AAG Val Cys Ala Ala Tyr Asp Ala Ile Leu Glu Arg Asn Val Ala Ile Lys 290 295 300	912
AAG CTA AGC CGA CCA TTT CAG AAT CAG ACT CAT GCC AAG CGG GCC TAC Lys Leu Ser Arg Pro Phe Gln Asn Gln Thr His Ala Lys Arg Ala Tyr 305 310 315 320	960
AGA GAG CTA GTT CTT ATG AAA TGT GTT AAT CAC AAA AAT ATA ATT GGC Arg Glu Leu Val Leu Met Lys Cys Val Asn His Lys Asn Ile Ile Gly 325 330 335	1008
CTT TTG AAT GTT TTC ACA CCA CAG AAA TCC CTA GAA GAA TTT CAA GAT Leu Leu Asn Val Phe Thr Pro Gln Lys Ser Leu Glu Glu Phe Gln Asp 340 345 350	1056
GTT TAC ATA GTC ATG GAG CTC ATG GAT GCA AAT CTT TGC CAA GTG ATT Val Tyr Ile Val Met Glu Leu Met Asp Ala Asn Leu Cys Gln Val Ile	1104

355	360	365	
CAG ATG GAG CTA GAT CAT GAA AGA ATG TCC TAC CTT CTC TAT CAG ATG Gln Met Glu Leu Asp His Glu Arg Met Ser Tyr Leu Leu Tyr Gln Met 370 375 380			1152
CTG TGT GGA ATC AAG CAC CTT CAT TCT GCT GGA ATT ATT CAT CGG GAC Leu Cys Gly Ile Lys His Leu His Ser Ala Gly Ile Ile His Arg Asp 385 390 395 400			1200
TTA AAG CCC AGT AAT ATA GTA GTA AAA TCT GAT TGC ACT TTG AAG ATT Leu Lys Pro Ser Asn Ile Val Val Lys Ser Asp Cys Thr Leu Lys Ile 405 410 415			1248
CTT GAC TTC GGT CTG GCC AGG ACT GCA GGA ACG AGT TTT ATG ATG ACG Leu Asp Phe Gly Leu Ala Arg Thr Ala Gly Thr Ser Phe Met Met Thr 420 425 430			1296
CCT TAT GTA GTG ACT CGC TAC TAC AGA GCA CCC GAG GTC ATC CTT GGC Pro Tyr Val Val Thr Arg Tyr Tyr Arg Ala Pro Glu Val Ile Leu Gly 435 440 445			1344
ATG GGC TAC AAG GAA AAC GTG GAT TTA TGG TCT GTG GGG TGC ATT ATG Met Gly Tyr Lys Glu Asn Val Asp Leu Trp Ser Val Gly Cys Ile Met 450 455 460			1392
GGA GAA ATG GTT TGC CAC AAA ATC CTC TTT CCA GGA AGG GAC TAT ATT Gly Glu Met Val Cys His Lys Ile Leu Phe Pro Gly Arg Asp Tyr Ile 465 470 475 480			1440
GAT CAG TGG AAT AAA GTT ATT GAA CAG CTT GGA ACA CCA TGT CCT GAA Asp Gln Trp Asn Lys Val Ile Glu Gln Leu Gly Thr Pro Cys Pro Glu 485 490 495			1488
TTC ATG AAG AAA CTG CAA CCA ACA GTA AGG ACT TAC GTT GAA AAC AGA Phe Met Lys Lys Leu Gln Pro Thr Val Arg Thr Tyr Val Glu Asn Arg 500 505 510			1536
CCT AAA TAT GCT GGA TAT AGC TTT GAG AAA CTC TTC CCT GAT GTC CTT Pro Lys Tyr Ala Gly Tyr Ser Phe Glu Lys Leu Phe Pro Asp Val Leu 515 520 525			1584
TTC CCA GCT GAC TCA GAA CAC AAC AAA CTT AAA GCC AGT CAG GCA AGG Phe Pro Ala Asp Ser Glu His Asn Lys Leu Lys Ala Ser Gln Ala Arg 530 535 540			1632
GAT TTG TTA TCC AAA ATG CTG GTA ATA GAT GCA TCT AAA AGG ATC TCT Asp Leu Leu Ser Lys Met Leu Val Ile Asp Ala Ser Lys Arg Ile Ser 545 550 555 560			1680
GTA GAT GAA GCT CTC CAA CAC CCG TAC ATC AAT GTC TGG TAT GAT CCT Val Asp Glu Ala Leu Gln His Pro Tyr Ile Asn Val Trp Tyr Asp Pro 565 570 575			1728
TCT GAA GCA GAA GCT CCA CCA CCA AAG ATC CCT GAC AAG CAG TTA GAT Ser Glu Ala Glu Ala Pro Pro Pro Lys Ile Pro Asp Lys Gln Leu Asp 580 585 590			1776

28

GAA AGG GAA CAC ACA ATA GAA GAG TGG AAA GAA TTG ATA TAT AAG GAA	1824
Glu Arg Glu His Thr Ile Glu Glu Trp Lys Glu Leu Ile Tyr Lys Glu	
595 600 605	
GTT ATG GAC TTG GAG GAG AGA ACC AAG AAT GGA GTT ATA CGG GGG CAG	1872
Val Met Asp Leu Glu Glu Arg Thr Lys Asn Gly Val Ile Arg Gly Gln	
610 615 620	
CCC TCT CCT TTA GCA CAG GTG CAG CAG TGA	1902
Pro Ser Pro Leu Ala Gln Val Gln Gln	
625 630	

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 633 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
100 105 110	
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly	
115 120 125	
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr	
130 135 140	
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn	
145 150 155 160	
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser	
165 170 175	
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly	
180 185 190	
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu	
195 200 205	
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe	
210 215 220	
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser	
225 230 235 240	
Gly Leu Arg Ser Arg Ala Arg Ala Ile Met Ser Arg Ser Lys Arg Asp	
245 250 255	

```

Asn Asn Phe Tyr Ser Val Glu Ile Gly Asp Ser Thr Phe Thr Val Leu
260                               265                               270
Lys Arg Tyr Gln Asn Leu Lys Pro Ile Gly Ser Gly Ala Gln Gly Ile
275                               280                               285
Val Cys Ala Ala Tyr Asp Ala Ile Leu Glu Arg Asn Val Ala Ile Lys
290                               295                               300
Lys Leu Ser Arg Pro Phe Gln Asn Gln Thr His Ala Lys Arg Ala Tyr
305                               310                               315                               320
Arg Glu Leu Val Leu Met Lys Cys Val Asn His Lys Asn Ile Ile Gly
325                               330                               335
Leu Leu Asn Val Phe Thr Pro Gln Lys Ser Leu Glu Glu Phe Gln Asp
340                               345                               350
Val Tyr Ile Val Met Glu Leu Met Asp Ala Asn Leu Cys Gln Val Ile
355                               360                               365
Gln Met Glu Leu Asp His Glu Arg Met Ser Tyr Leu Leu Tyr Gln Met
370                               375                               380
Leu Cys Gly Ile Lys His Leu His Ser Ala Gly Ile Ile His Arg Asp
385                               390                               395                               400
Leu Lys Pro Ser Asn Ile Val Val Lys Ser Asp Cys Thr Leu Lys Ile
405                               410                               415
Leu Asp Phe Gly Leu Ala Arg Thr Ala Gly Thr Ser Phe Met Met Thr
420                               425                               430
Pro Tyr Val Val Thr Arg Tyr Tyr Arg Ala Pro Glu Val Ile Leu Gly
435                               440                               445
Met Gly Tyr Lys Glu Asn Val Asp Leu Trp Ser Val Gly Cys Ile Met
450                               455                               460
Gly Glu Met Val Cys His Lys Ile Leu Phe Pro Gly Arg Asp Tyr Ile
465                               470                               475                               480
Asp Gln Trp Asn Lys Val Ile Glu Gln Leu Gly Thr Pro Cys Pro Glu
485                               490                               495
Phe Met Lys Lys Leu Gln Pro Thr Val Arg Thr Tyr Val Glu Asn Arg
500                               505                               510
Pro Lys Tyr Ala Gly Tyr Ser Phe Glu Lys Leu Phe Pro Asp Val Leu
515                               520                               525
Phe Pro Ala Asp Ser Glu His Asn Lys Leu Lys Ala Ser Gln Ala Arg
530                               535                               540
Asp Leu Leu Ser Lys Met Leu Val Ile Asp Ala Ser Lys Arg Ile Ser
545                               550                               555                               560
Val Asp Glu Ala Leu Gln His Pro Tyr Ile Asn Val Trp Tyr Asp Pro
565                               570                               575
Ser Glu Ala Glu Ala Pro Pro Pro Lys Ile Pro Asp Lys Gln Leu Asp
580                               585                               590
Glu Arg Glu His Thr Ile Glu Glu Trp Lys Glu Leu Ile Tyr Lys Glu
595                               600                               605
Val Met Asp Leu Glu Glu Arg Thr Lys Asn Gly Val Ile Arg Gly Gln
610                               615                               620
Pro Ser Pro Leu Ala Gln Val Gln Gln
625                               630

```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1824 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
 (B) LOCATION: 1...1821
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG	336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
100 105 110	
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC	384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly	
115 120 125	
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC	432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr	
130 135 140	
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC	480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn	
145 150 155 160	
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC	528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser	
165 170 175	
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC	576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly	
180 185 190	
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG	624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu	

195	200	205	
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC			672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe			
210	215	220	
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC			720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser			
225	230	235	240
GGA CTC AGA TCT CGA GGG AAA ATG TCT CAG GAG AGG CCC ACG TTC TAC			768
Gly Leu Arg Ser Arg Gly Lys Met Ser Gln Glu Arg Pro Thr Phe Tyr			
245	250	255	
CGG CAG GAG CTG AAC AAG ACA ATC TGG GAG GTG CCC GAG CGT TAC CAG			816
Arg Gln Glu Leu Asn Lys Thr Ile Trp Glu Val Pro Glu Arg Tyr Gln			
260	265	270	
AAC CTG TCT CCA GTG GGC TCT GGC GCC TAT GGC TCT GTG TGT GCT GCT			864
Asn Leu Ser Pro Val Gly Ser Gly Ala Tyr Gly Ser Val Cys Ala Ala			
275	280	285	
TTT GAC ACA AAA ACG GGG TTA CGT GTG GCA GTG AAG AAG CTC TCC AGA			912
Phe Asp Thr Lys Thr Gly Leu Arg Val Ala Val Lys Lys Leu Ser Arg			
290	295	300	
CCA TTT CAG TCC ATC ATT CAT GCG AAA AGA ACC TAC AGA GAA CTG CGG			960
Pro Phe Gln Ser Ile Ile His Ala Lys Arg Thr Tyr Arg Glu Leu Arg			
305	310	315	320
TTA CTT AAA CAT ATG AAA CAT GAA AAT GTG ATT GGT CTG TTG GAC GTT			1008
Leu Leu Lys His Met Lys His Glu Asn Val Ile Gly Leu Leu Asp Val			
325	330	335	
TTT ACA CCT GCA AGG TCT CTG GAG GAA TTC AAT GAT GTG TAT CTG GTG			1056
Phe Thr Pro Ala Arg Ser Leu Glu Glu Phe Asn Asp Val Tyr Leu Val			
340	345	350	
ACC CAT CTC ATG GGG GCA GAT CTG AAC AAC ATT GTG AAA TGT CAG AAG			1104
Thr His Leu Met Gly Ala Asp Leu Asn Asn Ile Val Lys Cys Gln Lys			
355	360	365	
CTT ACA GAT GAC CAT GTT CAG TTC CTT ATC TAC CAA ATT CTC CGA GGT			1152
Leu Thr Asp Asp His Val Gln Phe Leu Ile Tyr Gln Ile Leu Arg Gly			
370	375	380	
CTA AAG TAT ATA CAT TCA GCT GAC ATA ATT CAC AGG GAC CTA AAA CCT			1200
Leu Lys Tyr Ile His Ser Ala Asp Ile Ile His Arg Asp Leu Lys Pro			
385	390	395	400
AGT AAT CTA GCT GTG AAT GAA GAC TGT GAG CTG AAG ATT CTG GAT TTT			1248
Ser Asn Leu Ala Val Asn Glu Asp Cys Glu Leu Lys Ile Leu Asp Phe			
405	410	415	
GGA CTG GCT CGG CAC ACA GAT GAT GAA ATG ACA GGC TAC GTG GCC ACT			1296
Gly Leu Ala Arg His Thr Asp Asp Glu Met Thr Gly Tyr Val Ala Thr			
420	425	430	

52

AGG TGG TAC AGG GCT CCT GAG ATC ATG CTG AAC TGG ATG CAT TAC AAC Arg Trp Tyr Arg Ala Pro Glu Ile Met Leu Asn Trp Met His Tyr Asn 435 440 445	1344
CAG ACA GTT GAT ATT TGG TCA GTG GGA TGC ATA ATG GCC GAG CTG TTG Gln Thr Val Asp Ile Trp Ser Val Gly Cys Ile Met Ala Glu Leu Leu 450 455 460	1392
ACT GGA AGA ACA TTG TTT CCT GGT ACA GAC CAT ATT GAT CAG TTG AAG Thr Gly Arg Thr Leu Phe Pro Gly Thr Asp His Ile Asp Gln Leu Lys 465 470 475 480	1440
CTC ATT TTA AGA CTC GTT GGA ACC CCA GGG GCT GAG CTT TTG AAG AAA Leu Ile Leu Arg Leu Val Gly Thr Pro Gly Ala Glu Leu Leu Lys Lys 485 490 495	1488
ATC TCC TCA GAG TCT GCA AGA AAC TAT ATT CAG TCT TTG ACT CAG ATG Ile Ser Ser Glu Ser Ala Arg Asn Tyr Ile Gln Ser Leu Thr Gln Met 500 505 510	1536
CCG AAG ATG AAC TTT GCG AAT GTA TTT ATT GGT GCC AAT CCC CTG GCT Pro Lys Met Asn Phe Ala Asn Val Phe Ile Gly Ala Asn Pro Leu Ala 515 520 525	1584
GTC GAC TTG CTG GAG AAG ATG CTT GTA TTG GAC TCA GAT AAG AGA ATT Val Asp Leu Leu Glu Lys Met Leu Val Leu Asp Ser Asp Lys Arg Ile 530 535 540	1632
ACA GCG GCC CAA GCC CTT GCA CAT GCC TAC TTT GCT CAG TAC CAC GAT Thr Ala Ala Gln Ala Leu Ala His Ala Tyr Phe Ala Gln Tyr His Asp 545 550 555 560	1680
CCT GAT GAT GAA CCA GTG GCC GAT CCT TAT GAT CAG TCC TTT GAA AGC Pro Asp Asp Glu Pro Val Ala Asp Pro Tyr Asp Gln Ser Phe Glu Ser 565 570 575	1728
AGG GAC CTC CTT ATA GAT GAG TGG AAA AGC CTG ACC TAT GAT GAA GTC Arg Asp Leu Leu Ile Asp Glu Trp Lys Ser Leu Thr Tyr Asp Glu Val 580 585 590	1776
ATC AGC TTT GTG CCA CCA CCC CTT GAC CAA GAA GAG ATG GAG TCC TGA Ile Ser Phe Val Pro Pro Pro Leu Asp Gln Glu Glu Met Glu Ser 595 600 605	1824

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 607 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10 15
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 20 25 30
 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 35 40 45
 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50 55 60
 Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 65 70 75 80
 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85 90 95
 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 100 105 110
 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 115 120 125
 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 130 135 140
 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
 145 150 155 160
 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
 165 170 175
 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
 180 185 190
 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
 195 200 205
 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
 210 215 220
 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
 225 230 235 240
 Gly Leu Arg Ser Arg Gly Lys Met Ser Gln Glu Arg Pro Thr Phe Tyr
 245 250 255
 Arg Gln Glu Leu Asn Lys Thr Ile Trp Glu Val Pro Glu Arg Tyr Gln
 260 265 270
 Asn Leu Ser Pro Val Gly Ser Gly Ala Tyr Gly Ser Val Cys Ala Ala
 275 280 285
 Phe Asp Thr Lys Thr Gly Leu Arg Val Ala Val Lys Lys Leu Ser Arg
 290 295 300
 Pro Phe Gln Ser Ile Ile His Ala Lys Arg Thr Tyr Arg Glu Leu Arg
 305 310 315 320
 Leu Leu Lys His Met Lys His Glu Asn Val Ile Gly Leu Leu Asp Val
 325 330 335
 Phe Thr Pro Ala Arg Ser Leu Glu Glu Phe Asn Asp Val Tyr Leu Val
 340 345 350
 Thr His Leu Met Gly Ala Asp Leu Asn Asn Ile Val Lys Cys Gln Lys
 355 360 365
 Leu Thr Asp Asp His Val Gln Phe Leu Ile Tyr Gln Ile Leu Arg Gly
 370 375 380
 Leu Lys Tyr Ile His Ser Ala Asp Ile Ile His Arg Asp Leu Lys Pro
 385 390 395 400
 Ser Asn Leu Ala Val Asn Glu Asp Cys Glu Leu Lys Ile Leu Asp Phe
 405 410 415
 Gly Leu Ala Arg His Thr Asp Asp Glu Met Thr Gly Tyr Val Ala Thr
 420 425 430
 Arg Trp Tyr Arg Ala Pro Glu Ile Met Leu Asn Trp Met His Tyr Asn
 435 440 445
 Gln Thr Val Asp Ile Trp Ser Val Gly Cys Ile Met Ala Glu Leu Leu
 450 455 460

Thr Gly Arg Thr Leu Phe Pro Gly Thr Asp His Ile Asp Gln Leu Lys
 465 470 475 480
 Leu Ile Leu Arg Leu Val Gly Thr Pro Gly Ala Glu Leu Leu Lys Lys
 485 490 495
 Ile Ser Ser Glu Ser Ala Arg Asn Tyr Ile Gln Ser Leu Thr Gln Met
 500 505 510
 Pro Lys Met Asn Phe Ala Asn Val Phe Ile Gly Ala Asn Pro Leu Ala
 515 520 525
 Val Asp Leu Leu Glu Lys Met Leu Val Leu Asp Ser Asp Lys Arg Ile
 530 535 540
 Thr Ala Ala Gln Ala Leu Ala His Ala Tyr Phe Ala Gln Tyr His Asp
 545 550 555 560
 Pro Asp Asp Glu Pro Val Ala Asp Pro Tyr Asp Gln Ser Phe Glu Ser
 565 570 575
 Arg Asp Leu Leu Ile Asp Glu Trp Lys Ser Leu Thr Tyr Asp Glu Val
 580 585 590
 Ile Ser Phe Val Pro Pro Pro Leu Asp Gln Glu Glu Met Glu Ser
 595 600 605

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2907 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2904
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	

CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	336
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	
			100					105					110			
GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	CTG	AAG	GGC	384
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	
			115				120					125				
ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	CTG	GAG	TAC	432
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	
			130			135					140					
AAC	TAC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	480
Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	
145					150					155					160	
GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	528
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	
				165				170						175		
GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	576
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	
			180				185					190				
CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	TCC	GCC	CTG	624
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	
			195				200					205				
AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	CTG	GAG	TTC	672
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	
			210			215					220					
GTG	ACC	GCC	GCC	GGG	ATC	ACT	CTC	GGC	ATG	GAC	GAG	CTG	TAC	AAG	TCC	720
Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Ser	
225					230					235					240	
GGA	CTC	AGA	TCT	ATG	AGT	GCT	GAG	GGG	TAC	CAG	TAC	AGA	GCG	CTG	TAT	768
Gly	Leu	Arg	Ser	Met	Ser	Ala	Glu	Gly	Tyr	Gln	Tyr	Arg	Ala	Leu	Tyr	
				245				250					255			
GAT	TAT	AAA	AAG	GAA	AGA	GAA	GAA	GAT	ATT	GAC	TTG	CAC	TTG	GGT	GAC	816
Asp	Tyr	Lys	Lys	Glu	Arg	Glu	Glu	Asp	Ile	Asp	Leu	His	Leu	Gly	Asp	
			260				265					270				
ATA	TTG	ACT	GTG	AAT	AAA	GGG	TCC	TTA	GTA	GCT	CTT	GGA	TTC	AGT	GAT	864
Ile	Leu	Thr	Val	Asn	Lys	Gly	Ser	Leu	Val	Ala	Leu	Gly	Phe	Ser	Asp	
			275				280					285				
GGA	CAG	GAA	GCC	AGG	CCT	GAA	GAA	ATT	GGC	TGG	TTA	AAT	GGC	TAT	AAT	912
Gly	Gln	Glu	Ala	Arg	Pro	Glu	Glu	Ile	Gly	Trp	Leu	Asn	Gly	Tyr	Asn	
			290			295				300						
GAA	ACC	ACA	GGG	GAA	AGG	GGG	GAC	TTT	CCG	GGA	ACT	TAC	GTA	GAA	TAT	

ATT GGA AGG AAA AAA ATC TCG CCT CCC ACA CCA AAG CCC CGG CCA CCT Ile Gly Arg Lys Lys Ile Ser Pro Pro Thr Pro Lys Pro Arg Pro Pro 325 330 335	1008
CGG CCT CTT CCT GTT GCA CCA GGT TCT TCG AAA ACT GAA GCA GAT GTT Arg Pro Leu Pro Val Ala Pro Gly Ser Ser Lys Thr Glu Ala Asp Val 340 345 350	1056
GAA CAA CAA GCT TTG ACT CTC CCG GAT CTT GCA GAG CAG TTT GCC CCT Glu Gln Gln Ala Leu Thr Leu Pro Asp Leu Ala Glu Gln Phe Ala Pro 355 360 365	1104
CCT GAC ATT GCC CCG CCT CTT CTT ATC AAG CTC GTG GAA GCC ATT GAA Pro Asp Ile Ala Pro Pro Leu Leu Ile Lys Leu Val Glu Ala Ile Glu 370 375 380	1152
AAG AAA GGT CTG GAA TGT TCA ACT CTA TAC AGA ACA CAG AGC TCC AGC Lys Lys Gly Leu Glu Cys Ser Thr Leu Tyr Arg Thr Gln Ser Ser Ser 385 390 395 400	1200
AAC CTG GCA GAA TTA CGA CAG CTT CTT GAT TGT GAT ACA CCC TCC GTG Asn Leu Ala Glu Leu Arg Gln Leu Leu Asp Cys Asp Thr Pro Ser Val 405 410 415	1248
GAC TTG GAA ATG ATC GAT GTG CAC GTT TTG GCT GAC GCT TTC AAA CGC Asp Leu Glu Met Ile Asp Val His Val Leu Ala Asp Ala Phe Lys Arg 420 425 430	1296
TAT CTC CTG GAC TTA CCA AAT CCT GTC ATT CCA GCA GCC GTT TAC AGT Tyr Leu Leu Asp Leu Pro Asn Pro Val Ile Pro Ala Ala Val Tyr Ser 435 440 445	1344
GAA ATG ATT TCT TTA GCT CCA GAA GTA CAA AGC TCC GAA GAA TAT ATT Glu Met Ile Ser Leu Ala Pro Glu Val Gln Ser Ser Glu Glu Tyr Ile 450 455 460	1392
CAG CTA TTG AAG AAG CTT ATT AGG TCG CCT AGC ATA CCT CAT CAG TAT Gln Leu Leu Lys Lys Leu Ile Arg Ser Pro Ser Ile Pro His Gln Tyr 465 470 475 480	1440
TGG CTT ACG CTT CAG TAT TTG TTA AAA CAT TTC TTC AAG CTC TCT CAA Trp Leu Thr Leu Gln Tyr Leu Leu Lys His Phe Phe Lys Leu Ser Gln 485 490 495	1488
ACC TCC AGC AAA AAT CTG TTG AAT GCA AGA GTA CTC TCT GAA ATT TTC Thr Ser Ser Lys Asn Leu Leu Asn Ala Arg Val Leu Ser Glu Ile Phe 500 505 510	1536
AGC CCT ATG CTT TTC AGA TTC TCA GCA GCC AGC TCT GAT AAT ACT GAA Ser Pro Met Leu Phe Arg Phe Ser Ala Ala Ser Ser Asp Asn Thr Glu 515 520 525	1584
AAC CTC ATA AAA GTT ATA GAA ATT TTA ATC TCA ACT GAA TGG AAT GAA Asn Leu Ile Lys Val Ile Glu Ile Leu Ile Ser Thr Glu Trp Asn Glu 530 535 540	1632
CGA CAG CCT GCA CCA GCA CTG CCT CCT AAA CCA CCA AAA CCT ACT ACT Arg Gln Pro Ala Pro Ala Leu Pro Pro Lys Pro Pro Lys Pro Thr Thr	1680

545	550	555	560	
GTA GCC AAC AAC GGT ATG AAT AAC AAT ATG TCC TTA CAA AAT GCT GAA				1728
Val Ala Asn Asn Gly Met Asn Asn Asn Met Ser Leu Gln Asn Ala Glu				
565	570	575		
TGG TAC TGG GGA GAT ATC TCG AGG GAA GAA GTG AAT GAA AAA CTT CGA				1776
Trp Tyr Trp Gly Asp Ile Ser Arg Glu Glu Val Asn Glu Lys Leu Arg				
580	585	590		
GAT ACA GCA GAC GGG ACC TTT TTG GTA CGA GAT GCG TCT ACT AAA ATG				1824
Asp Thr Ala Asp Gly Thr Phe Leu Val Arg Asp Ala Ser Thr Lys Met				
595	600	605		
CAT GGT GAT TAT ACT CTT ACA CTA AGG AAA GGG GGA AAT AAC AAA TTA				1872
His Gly Asp Tyr Thr Leu Thr Leu Arg Lys Gly Gly Asn Asn Lys Leu				
610	615	620		
ATC AAA ATA TTT CAT CGA GAT GGG AAA TAT GGC TTC TCT GAC CCA TTA				1920
Ile Lys Ile Phe His Arg Asp Gly Lys Tyr Gly Phe Ser Asp Pro Leu				
625	630	635	640	
ACC TTC AGT TCT GTG GTT GAA TTA ATA AAC CAC TAC CGG AAT GAA TCT				1968
Thr Phe Ser Ser Val Val Glu Leu Ile Asn His Tyr Arg Asn Glu Ser				
645	650	655		
CTA GCT CAG TAT AAT CCC AAA TTG GAT GTG AAA TTA CTT TAT CCA GTA				2016
Leu Ala Gln Tyr Asn Pro Lys Leu Asp Val Lys Leu Leu Tyr Pro Val				
660	665	670		
TCC AAA TAC CAA CAG GAT CAA GTT GTC AAA GAA GAT AAT ATT GAA GCT				2064
Ser Lys Tyr Gln Gln Asp Gln Val Val Lys Glu Asp Asn Ile Glu Ala				
675	680	685		
GTA GGG AAA AAA TTA CAT GAA TAT AAC ACT CAG TTT CAA GAA AAA AGT				2112
Val Gly Lys Lys Leu His Glu Tyr Asn Thr Gln Phe Gln Glu Lys Ser				
690	695	700		
CGA GAA TAT GAT AGA TTA TAT GAA GAA TAT ACC CGC ACA TCC CAG GAA				2160
Arg Glu Tyr Asp Arg Leu Tyr Glu Glu Tyr Thr Arg Thr Ser Gln Glu				
705	710	715	720	
ATC CAA ATG AAA AGG ACA GCT ATT GAA GCA TTT AAT GAA ACC ATA AAA				2208
Ile Gln Met Lys Arg Thr Ala Ile Glu Ala Phe Asn Glu Thr Ile Lys				
725	730	735		
ATA TTT GAA GAA CAG TGC CAG ACC CAA GAG CGG TAC AGC AAA GAA TAC				2256
Ile Phe Glu Glu Gln Cys Gln Thr Gln Glu Arg Tyr Ser Lys Glu Tyr				
740	745	750		
ATA GAA AAG TTT AAA CGT GAA GGC AAT GAG AAA GAA ATA CAA AGG ATT				2304
Ile Glu Lys Phe Lys Arg Glu Gly Asn Glu Lys Glu Ile Gln Arg Ile				
755	760	765		
ATG CAT AAT TAT GAT AAG TTG AAG TCT CGA ATC AGT GAA ATT ATT GAC				2352
Met His Asn Tyr Asp Lys Leu Lys Ser Arg Ile Ser Glu Ile Ile Asp				
770	775	780		

AGT AGA AGA AGA TTG GAA GAA GAC TTG AAG AAG CAG GCA GCT GAG TAT	2400
Ser Arg Arg Arg Leu Glu Glu Asp Leu Lys Lys Gln Ala Ala Glu Tyr	
785 790 795 800	
CGA GAA ATT GAC AAA CGT ATG AAC AGC ATT AAA CCA GAC CTT ATC CAG	2448
Arg Glu Ile Asp Lys Arg Met Asn Ser Ile Lys Pro Asp Leu Ile Gln	
805 810 815	
CTG AGA AAG ACG AGA GAC CAA TAC TTG ATG TGG TTG ACT CAA AAA GGT	2496
Leu Arg Lys Thr Arg Asp Gln Tyr Leu Met Trp Leu Thr Gln Lys Gly	
820 825 830	
GTT CGG CAA AAG AAG TTG AAC GAG TGG TTG GGC AAT GAA AAC ACT GAA	2544
Val Arg Gln Lys Lys Leu Asn Glu Trp Leu Gly Asn Glu Asn Thr Glu	
835 840 845	
GAC CAA TAT TCA CTG GTG GAA GAT GAT GAA GAT TTG CCC CAT CAT GAT	2592
Asp Gln Tyr Ser Leu Val Glu Asp Asp Glu Asp Leu Pro His His Asp	
850 855 860	
GAG AAG ACA TGG AAT GTT GGA AGC AGC AAC CGA AAC AAA GCT GAA AAC	2640
Glu Lys Thr Trp Asn Val Gly Ser Ser Asn Arg Asn Lys Ala Glu Asn	
865 870 875 880	
CTG TTG CGA GGG AAG CGA GAT GGC ACT TTT CTT GTC CGG GAG AGC AGT	2688
Leu Leu Arg Gly Lys Arg Asp Gly Thr Phe Leu Val Arg Glu Ser Ser	
885 890 895	
AAA CAG GGC TGC TAT GCC TGC TCT GTA GTG GTG GAC GGC GAA GTA AAG	2736
Lys Gln Gly Cys Tyr Ala Cys Ser Val Val Val Asp Gly Glu Val Lys	
900 905 910	
CAT TGT GTC ATA AAC AAA ACA GCA ACT GGC TAT GGC TTT GCC GAG CCC	2784
His Cys Val Ile Asn Lys Thr Ala Thr Gly Tyr Gly Phe Ala Glu Pro	
915 920 925	
TAT AAC TTG TAC AGC TCT CTG AAA GAA CTG GTG CTA CAT TAC CAA CAC	2832
Tyr Asn Leu Tyr Ser Ser Leu Lys Glu Leu Val Leu His Tyr Gln His	
930 935 940	
ACC TCC CTT GTG CAG CAC AAC GAC TCC CTC AAT GTC ACA CTA GCC TAC	2880
Thr Ser Leu Val Gln His Asn Asp Ser Leu Asn Val Thr Leu Ala Tyr	
945 950 955 960	
CCA GTA TAT GCA CAG CAG AGG CGA TGA	2907
Pro Val Tyr Ala Gln Gln Arg Arg	
965	

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 968 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1           5           10           15
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
           20           25           30
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 35           40           45
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50           55           60
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 65           70           75           80
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
           85           90           95
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
           100          105          110
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
           115          120          125
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 130          135          140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
 145          150          155          160
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
           165          170          175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
           180          185          190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
           195          200          205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
 210          215          220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
 225          230          235          240
Gly Leu Arg Ser Met Ser Ala Glu Gly Tyr Gln Tyr Arg Ala Leu Tyr
           245          250          255
Asp Tyr Lys Lys Glu Arg Glu Glu Asp Ile Asp Leu His Leu Gly Asp
           260          265          270
Ile Leu Thr Val Asn Lys Gly Ser Leu Val Ala Leu Gly Phe Ser Asp
           275          280          285
Gly Gln Glu Ala Arg Pro Glu Glu Ile Gly Trp Leu Asn Gly Tyr Asn
 290          295          300
Glu Thr Thr Gly Glu Arg Gly Asp Phe Pro Gly Thr Tyr Val Glu Tyr
 305          310          315          320
Ile Gly Arg Lys Lys Ile Ser Pro Pro Thr Pro Lys Pro Arg Pro Pro
           325          330          335
Arg Pro Leu Pro Val Ala Pro Gly Ser Ser Lys Thr Glu Ala Asp Val
           340          345          350
Glu Gln Gln Ala Leu Thr Leu Pro Asp Leu Ala Glu Gln Phe Ala Pro
           355          360          365
Pro Asp Ile Ala Pro Pro Leu Leu Ile Lys Leu Val Glu Ala Ile Glu
 370          375          380
Lys Lys Gly Leu Glu Cys Ser Thr Leu Tyr Arg Thr Gln Ser Ser Ser
 385          390          395          400
Asn Leu Ala Glu Leu Arg Gln Leu Leu Asp Cys Asp Thr Pro Ser Val
           405          410          415
Asp Leu Glu Met Ile Asp Val His Val Leu Ala Asp Ala Phe Lys Arg
           420          425          430

```

Tyr Leu Leu Asp Leu Pro Asn Pro Val Ile Pro Ala Ala Val Tyr Ser
 435 440 445
 Glu Met Ile Ser Leu Ala Pro Glu Val Gln Ser Ser Glu Glu Tyr Ile
 450 455 460
 Gln Leu Leu Lys Lys Leu Ile Arg Ser Pro Ser Ile Pro His Gln Tyr
 465 470 475 480
 Trp Leu Thr Leu Gln Tyr Leu Leu Lys His Phe Phe Lys Leu Ser Gln
 485 490 495
 Thr Ser Ser Lys Asn Leu Leu Asn Ala Arg Val Leu Ser Glu Ile Phe
 500 505 510
 Ser Pro Met Leu Phe Arg Phe Ser Ala Ala Ser Ser Asp Asn Thr Glu
 515 520 525
 Asn Leu Ile Lys Val Ile Glu Ile Leu Ile Ser Thr Glu Trp Asn Glu
 530 535 540
 Arg Gln Pro Ala Pro Ala Leu Pro Pro Lys Pro Pro Lys Pro Thr Thr
 545 550 555 560
 Val Ala Asn Asn Gly Met Asn Asn Asn Met Ser Leu Gln Asn Ala Glu
 565 570 575
 Trp Tyr Trp Gly Asp Ile Ser Arg Glu Glu Val Asn Glu Lys Leu Arg
 580 585 590
 Asp Thr Ala Asp Gly Thr Phe Leu Val Arg Asp Ala Ser Thr Lys Met
 595 600 605
 His Gly Asp Tyr Thr Leu Thr Leu Arg Lys Gly Gly Asn Asn Lys Leu
 610 615 620
 Ile Lys Ile Phe His Arg Asp Gly Lys Tyr Gly Phe Ser Asp Pro Leu
 625 630 635 640
 Thr Phe Ser Ser Val Val Glu Leu Ile Asn His Tyr Arg Asn Glu Ser
 645 650 655
 Leu Ala Gln Tyr Asn Pro Lys Leu Asp Val Lys Leu Leu Tyr Pro Val
 660 665 670
 Ser Lys Tyr Gln Gln Asp Gln Val Val Lys Glu Asp Asn Ile Glu Ala
 675 680 685
 Val Gly Lys Lys Leu His Glu Tyr Asn Thr Gln Phe Gln Glu Lys Ser
 690 695 700
 Arg Glu Tyr Asp Arg Leu Tyr Glu Glu Tyr Thr Arg Thr Ser Gln Glu
 705 710 715 720
 Ile Gln Met Lys Arg Thr Ala Ile Glu Ala Phe Asn Glu Thr Ile Lys
 725 730 735
 Ile Phe Glu Glu Gln Cys Gln Thr Gln Glu Arg Tyr Ser Lys Glu Tyr
 740 745 750
 Ile Glu Lys Phe Lys Arg Glu Gly Asn Glu Lys Glu Ile Gln Arg Ile
 755 760 765
 Met His Asn Tyr Asp Lys Leu Lys Ser Arg Ile Ser Glu Ile Ile Asp
 770 775 780
 Ser Arg Arg Arg Leu Glu Asp Leu Lys Lys Gln Ala Ala Glu Tyr
 785 790 795 800
 Arg Glu Ile Asp Lys Arg Met Asn Ser Ile Lys Pro Asp Leu Ile Gln
 805 810 815
 Leu Arg Lys Thr Arg Asp Gln Tyr Leu Met Trp Leu Thr Gln Lys Gly
 820 825 830
 Val Arg Gln Lys Lys Leu Asn Glu Trp Leu Gly Asn Glu Asn Thr Glu
 835 840 845
 Asp Gln Tyr Ser Leu Val Glu Asp Asp Glu Asp Leu Pro His His Asp
 850 855 860
 Glu Lys Thr Trp Asn Val Gly Ser Ser Asn Arg Asn Lys Ala Glu Asn
 865 870 875 880
 Leu Leu Arg Gly Lys Arg Asp Gly Thr Phe Leu Val Arg Glu Ser Ser
 885 890 895

Lys Gln Gly Cys Tyr Ala Cys Ser Val Val Val Asp Gly Glu Val Lys
 900 905 910
 His Cys Val Ile Asn Lys Thr Ala Thr Gly Tyr Gly Phe Ala Glu Pro
 915 920 925
 Tyr Asn Leu Tyr Ser Ser Leu Lys Glu Leu Val Leu His Tyr Gln His
 930 935 940
 Thr Ser Leu Val Gln His Asn Asp Ser Leu Asn Val Thr Leu Ala Tyr
 945 950 955 960
 Pro Val Tyr Ala Gln Gln Arg Arg
 965

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2160 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2157
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG	336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
100 105 110	
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC	384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly	

72

115	120	125	
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC			432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr			
130	135	140	
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC			480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn			
145	150	155	160
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC			528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser			
165	170	175	
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC			576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly			
180	185	190	
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG			624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu			
195	200	205	
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC			672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe			
210	215	220	
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC			720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser			
225	230	235	240
GGA CTC AGA TCT CGA GCT CAA GCT TCG AAT TCG ACC ATG TCG TCC ATC			768
Gly Leu Arg Ser Arg Ala Gln Ala Ser Asn Ser Thr Met Ser Ser Ile			
245	250	255	
TTG CCA TTC ACG CCG CCA GTT GTG AAG AGA CTG CTG GGA TGG AAG AAG			816
Leu Pro Phe Thr Pro Pro Val Val Lys Arg Leu Leu Gly Trp Lys Lys			
260	265	270	
TCA GCT GGT GGG TCT GGA GGA GCA GGC GGA GGA GAG CAG AAT GGG CAG			864
Ser Ala Gly Gly Ser Gly Gly Ala Gly Gly Gly Glu Gln Asn Gly Gln			
275	280	285	
GAA GAA AAG TGG TGT GAG AAA GCA GTG AAA AGT CTG GTG AAG AAG CTA			912
Glu Glu Lys Trp Cys Glu Lys Ala Val Lys Ser Leu Val Lys Lys Leu			
290	295	300	
AAG AAA ACA GGA CGA TTA GAT GAG CTT GAG AAA GCC ATC ACC ACT CAA			960
Lys Lys Thr Gly Arg Leu Asp Glu Leu Glu Lys Ala Ile Thr Thr Gln			
305	310	315	320
AAC TGT AAT ACT AAA TGT GTT ACC ATA CCA AGC ACT TGC TCT GAA ATT			1008
Asn Cys Asn Thr Lys Cys Val Thr Ile Pro Ser Thr Cys Ser Glu Ile			
325	330	335	
TGG GGA CTG AGT ACA CCA AAT ACG ATA GAT CAG TGG GAT ACA ACA GGC			1056
Trp Gly Leu Ser Thr Pro Asn Thr Ile Asp Gln Trp Asp Thr Thr Gly			
340	345	350	

CTT TAC AGC TTC TCT GAA CAA ACC AGG TCT CTT GAT GGT CGT CTC CAG Leu Tyr Ser Phe Ser Glu Gln Thr Arg Ser Leu Asp Gly Arg Leu Gln 355 360 365	1104
GTA TCC CAT CGA AAA GGA TTG CCA CAT GTT ATA TAT TGC CGA TTA TGG Val Ser His Arg Lys Gly Leu Pro His Val Ile Tyr Cys Arg Leu Trp 370 375 380	1152
CGC TGG CCT GAT CTT CAC AGT CAT CAT GAA CTC AAG GCA ATT GAA AAC Arg Trp Pro Asp Leu His Ser His His Glu Leu Lys Ala Ile Glu Asn 385 390 395 400	1200
TGC GAA TAT GCT TTT AAT CTT AAA AAG GAT GAA GTA TGT GTA AAC CCT Cys Glu Tyr Ala Phe Asn Leu Lys Lys Asp Glu Val Cys Val Asn Pro 405 410 415	1248
TAC CAC TAT CAG AGA GTT GAG ACA CCA GTT TTG CCT CCA GTA TTA GTG Tyr His Tyr Gln Arg Val Glu Thr Pro Val Leu Pro Pro Val Leu Val 420 425 430	1296
CCC CGA CAC ACC GAG ATC CTA ACA GAA CTT CCG CCT CTG GAT GAC TAT Pro Arg His Thr Glu Ile Leu Thr Glu Leu Pro Pro Leu Asp Asp Tyr 435 440 445	1344
ACT CAC TCC ATT CCA GAA AAC ACT AAC TTC CCA GCA GGA ATT GAG CCA Thr His Ser Ile Pro Glu Asn Thr Asn Phe Pro Ala Gly Ile Glu Pro 450 455 460	1392
CAG AGT AAT TAT ATT CCA GAA ACG CCA CCT CCT GGA TAT ATC AGT GAA Gln Ser Asn Tyr Ile Pro Glu Thr Pro Pro Pro Gly Tyr Ile Ser Glu 465 470 475 480	1440
GAT GGA GAA ACA AGT GAC CAA CAG TTG AAT CAA AGT ATG GAC ACA GGC Asp Gly Glu Thr Ser Asp Gln Gln Leu Asn Gln Ser Met Asp Thr Gly 485 490 495	1488
TCT CCA GCA GAA CTA TCT CCT ACT ACT CTT TCC CCT GTT AAT CAT AGC Ser Pro Ala Glu Leu Ser Pro Thr Thr Leu Ser Pro Val Asn His Ser 500 505 510	1536
TTG GAT TTA CAG CCA GTT ACT TAC TCA GAA CCT GCA TTT TGG TGT TCA Leu Asp Leu Gln Pro Val Thr Tyr Ser Glu Pro Ala Phe Trp Cys Ser 515 520 525	1584
ATA GCA TAT TAT GAA TTA AAT CAG AGG GTT GGA GAA ACC TTC CAT GCA Ile Ala Tyr Tyr Glu Leu Asn Gln Arg Val Gly Glu Thr Phe His Ala 530 535 540	1632
TCA CAG CCC TCA CTC ACT GTA GAT GGC TTT ACA GAC CCA TCA AAT TCA Ser Gln Pro Ser Leu Thr Val Asp Gly Phe Thr Asp Pro Ser Asn Ser 545 550 555 560	1680
GAG AGG TTC TGC TTA GGT TTA CTC TCC AAT GTT AAC CGA AAT GCC ACG Glu Arg Phe Cys Leu Gly Leu Leu Ser Asn Val Asn Arg Asn Ala Thr 565 570 575	1728
GTA GAA ATG ACA AGA AGG CAT ATA GGA AGA GGA GTG CGC TTA TAC TAC Val Glu Met Thr Arg Arg His Ile Gly Arg Gly Val Arg Leu Tyr Tyr 580 585 590 595	1776

777

580	585	590	
ATA GGT GGG GAA GTT TTT GCT GAG TGC CTA AGT GAT AGT GCA ATC TTT			1824
Ile Gly Gly Glu Val Phe Ala Glu Cys Leu Ser Asp Ser Ala Ile Phe			
595	600	605	
GTG CAG AGC CCC AAT TGT AAT CAG AGA TAT GGC TGG CAC CCT GCA ACA			1872
Val Gln Ser Pro Asn Cys Asn Gln Arg Tyr Gly Trp His Pro Ala Thr			
610	615	620	
GTG TGT AAA ATT CCA CCA GGC TGT AAT CTG AAG ATC TTC AAC AAC CAG			1920
Val Cys Lys Ile Pro Pro Gly Cys Asn Leu Lys Ile Phe Asn Asn Gln			
625	630	635	640
GAA TTT GCT GCT CTT CTG GCT CAG TCT GTT AAT CAG GGT TTT GAA GCC			1968
Glu Phe Ala Ala Leu Leu Ala Gln Ser Val Asn Gln Gly Phe Glu Ala			
645	650	655	
GTC TAT CAG CTA ACT AGA ATG TGC ACC ATA AGA ATG AGT TTT GTG AAA			2016
Val Tyr Gln Leu Thr Arg Met Cys Thr Ile Arg Met Ser Phe Val Lys			
660	665	670	
GGG TGG GGA GCA GAA TAC CGA AGG CAG ACG GTA ACA AGT ACT CCT TGC			2064
Gly Trp Gly Ala Glu Tyr Arg Arg Gln Thr Val Thr Ser Thr Pro Cys			
675	680	685	
TGG ATT GAA CTT CAT CTG AAT GGA CCT CTA CAG TGG TTG GAC AAA GTA			2112
Trp Ile Glu Leu His Leu Asn Gly Pro Leu Gln Trp Leu Asp Lys Val			
690	695	700	
TTA ACT CAG ATG GGA TCC CCT TCA GTG CGT TGC TCA AGC ATG TCA TAA			2160
Leu Thr Gln Met Gly Ser Pro Ser Val Arg Cys Ser Ser Met Ser			
705	710	715	

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 719 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu			
1	5	10	15
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly			
20	25	30	
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile			
35	40	45	
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr			
50	55	60	
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys			
65	70	75	80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85 90 95
 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 100 105 110
 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 115 120 125
 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 130 135 140
 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
 145 150 155 160
 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
 165 170 175
 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
 180 185 190
 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
 195 200 205
 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
 210 215 220
 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
 225 230 235 240
 Gly Leu Arg Ser Arg Ala Gln Ala Ser Asn Ser Thr Met Ser Ser Ile
 245 250 255
 Leu Pro Phe Thr Pro Pro Val Val Lys Arg Leu Leu Gly Trp Lys Lys
 260 265 270
 Ser Ala Gly Gly Ser Gly Gly Ala Gly Gly Gly Glu Gln Asn Gly Gln
 275 280 285
 Glu Glu Lys Trp Cys Glu Lys Ala Val Lys Ser Leu Val Lys Lys Leu
 290 295 300
 Lys Lys Thr Gly Arg Leu Asp Glu Leu Glu Lys Ala Ile Thr Thr Gln
 305 310 315 320
 Asn Cys Asn Thr Lys Cys Val Thr Ile Pro Ser Thr Cys Ser Glu Ile
 325 330 335
 Trp Gly Leu Ser Thr Pro Asn Thr Ile Asp Gln Trp Asp Thr Thr Gly
 340 345 350
 Leu Tyr Ser Phe Ser Glu Gln Thr Arg Ser Leu Asp Gly Arg Leu Gln
 355 360 365
 Val Ser His Arg Lys Gly Leu Pro His Val Ile Tyr Cys Arg Leu Trp
 370 375 380
 Arg Trp Pro Asp Leu His Ser His His Glu Leu Lys Ala Ile Glu Asn
 385 390 395 400
 Cys Glu Tyr Ala Phe Asn Leu Lys Lys Asp Glu Val Cys Val Asn Pro
 405 410 415
 Tyr His Tyr Gln Arg Val Glu Thr Pro Val Leu Pro Pro Val Leu Val
 420 425 430
 Pro Arg His Thr Glu Ile Leu Thr Glu Leu Pro Pro Leu Asp Asp Tyr
 435 440 445
 Thr His Ser Ile Pro Glu Asn Thr Asn Phe Pro Ala Gly Ile Glu Pro
 450 455 460
 Gln Ser Asn Tyr Ile Pro Glu Thr Pro Pro Pro Gly Tyr Ile Ser Glu
 465 470 475 480
 Asp Gly Glu Thr Ser Asp Gln Gln Leu Asn Gln Ser Met Asp Thr Gly
 485 490 495
 Ser Pro Ala Glu Leu Ser Pro Thr Thr Leu Ser Pro Val Asn His Ser
 500 505 510
 Leu Asp Leu Gln Pro Val Thr Tyr Ser Glu Pro Ala Phe Trp Cys Ser
 515 520 525
 Ile Ala Tyr Tyr Glu Leu Asn Gln Arg Val Gly Glu Thr Phe His Ala
 530 535 540

72

```

Ser Gln Pro Ser Leu Thr Val Asp Gly Phe Thr Asp Pro Ser Asn Ser
545                      550                      555                      560
Glu Arg Phe Cys Leu Gly Leu Leu Ser Asn Val Asn Arg Asn Ala Thr
                      565                      570                      575
Val Glu Met Thr Arg Arg His Ile Gly Arg Gly Val Arg Leu Tyr Tyr
                      580                      585                      590
Ile Gly Gly Glu Val Phe Ala Glu Cys Leu Ser Asp Ser Ala Ile Phe
                      595                      600                      605
Val Gln Ser Pro Asn Cys Asn Gln Arg Tyr Gly Trp His Pro Ala Thr
                      610                      615                      620
Val Cys Lys Ile Pro Pro Gly Cys Asn Leu Lys Ile Phe Asn Asn Gln
625                      630                      635                      640
Glu Phe Ala Ala Leu Leu Ala Gln Ser Val Asn Gln Gly Phe Glu Ala
                      645                      650                      655
Val Tyr Gln Leu Thr Arg Met Cys Thr Ile Arg Met Ser Phe Val Lys
                      660                      665                      670
Gly Trp Gly Ala Glu Tyr Arg Arg Gln Thr Val Thr Ser Thr Pro Cys
                      675                      680                      685
Trp Ile Glu Leu His Leu Asn Gly Pro Leu Gln Trp Leu Asp Lys Val
690                      695                      700
Leu Thr Gln Met Gly Ser Pro Ser Val Arg Cys Ser Ser Met Ser
705                      710                      715

```

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2421 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2418
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG      48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
  1                      5                      10                      15

GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC      96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
                20                      25                      30

GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC     144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
                35                      40                      45

TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC     192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
                50                      55                      60

CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG     240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys

```

65	70	75	80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	85	90	95	288
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	100	105	110	336
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly	115	120	125	384
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr	130	135	140	432
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn	145	150	155	480
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser	165	170	175	528
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly	180	185	190	576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu	195	200	205	624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe	210	215	220	672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser	225	230	235	720
GGA CTC AGA TCT CGA GCT CAA GCT TCG AAT TCG AAT TCA ACC ATG GAC Gly Leu Arg Ser Arg Ala Gln Ala Ser Asn Ser Asn Ser Thr Met Asp	245	250	255	768
AAT ATG TCT ATT ACG AAT ACA CCA ACA AGT AAT GAT GCC TGT CTG AGC Asn Met Ser Ile Thr Asn Thr Pro Thr Ser Asn Asp Ala Cys Leu Ser	260	265	270	816
ATT GTG CAT AGT TTG ATG TGC CAT AGA CAA GGT GGA GAG AGT GAA ACA Ile Val His Ser Leu Met Cys His Arg Gln Gly Gly Glu Ser Glu Thr	275	280	285	864
TTT GCA AAA AGA GCA ATT GAA AGT TTG GTA AAG AAG CTG AAG GAG AAA Phe Ala Lys Arg Ala Ile Glu Ser Leu Val Lys Lys Leu Lys Glu Lys	290	295	300	912

AAA GAT GAA TTG GAT TCT TTA ATA ACA GCT ATA ACT ACA AAT GGA GCT 960
 Lys Asp Glu Leu Asp Ser Leu Ile Thr Ala Ile Thr Thr Asn Gly Ala
 305 310 315 320

CAT CCT AGT AAA TGT GTT ACC ATA CAG AGA ACA TTG GAT GGG AGG CTT 1008
 His Pro Ser Lys Cys Val Thr Ile Gln Arg Thr Leu Asp Gly Arg Leu
 325 330 335

CAG GTG GCT GGT CGG AAA GGA TTT CCT CAT GTG ATC TAT GCC CGT CTC 1056
 Gln Val Ala Gly Arg Lys Gly Phe Pro His Val Ile Tyr Ala Arg Leu
 340 345 350

TGG AGG TGG CCT GAT CTT CAC AAA AAT GAA CTA AAA CAT GTT AAA TAT 1104
 Trp Arg Trp Pro Asp Leu His Lys Asn Glu Leu Lys His Val Lys Tyr
 355 360 365

TGT CAG TAT GCG TTT GAC TTA AAA TGT GAT AGT GTC TGT GTG AAT CCA 1152
 Cys Gln Tyr Ala Phe Asp Leu Lys Cys Asp Ser Val Cys Val Asn Pro
 370 375 380

TAT CAC TAC GAA CGA GTT GTA TCA CCT GGA ATT GAT CTC TCA GGA TTA 1200
 Tyr His Tyr Glu Arg Val Val Ser Pro Gly Ile Asp Leu Ser Gly Leu
 385 390 395 400

ACA CTG CAG AGT AAT GCT CCA TCA AGT ATG ATG GTG AAG GAT GAA TAT 1248
 Thr Leu Gln Ser Asn Ala Pro Ser Ser Met Met Val Lys Asp Glu Tyr
 405 410 415

GTG CAT GAC TTT GAG GGA CAG CCA TCG TTG TCC ACT GAA GGA CAT TCA 1296
 Val His Asp Phe Glu Gly Gln Pro Ser Leu Ser Thr Glu Gly His Ser
 420 425 430

ATT CAA ACC ATC CAG CAT CCA CCA AGT AAT CGT GCA TCG ACA GAG ACA 1344
 Ile Gln Thr Ile Gln His Pro Pro Ser Asn Arg Ala Ser Thr Glu Thr
 435 440 445

TAC AGC ACC CCA GCT CTG TTA GCC CCA TCT GAG TCT AAT GCT ACC AGC 1392
 Tyr Ser Thr Pro Ala Leu Leu Ala Pro Ser Glu Ser Asn Ala Thr Ser
 450 455 460

ACT GCC AAC TTT CCC AAC ATT CCT GTG GCT TCC ACA AGT CAG CCT GCC 1440
 Thr Ala Asn Phe Pro Asn Ile Pro Val Ala Ser Thr Ser Gln Pro Ala
 465 470 475 480

AGT ATA CTG GGG GGC AGC CAT AGT GAA GGA CTG TTG CAG ATA GCA TCA 1488
 Ser Ile Leu Gly Gly Ser His Ser Glu Gly Leu Leu Gln Ile Ala Ser
 485 490 495

GGG CCT CAG CCA GGA CAG CAG CAG AAT GGA TTT ACT GGT CAG CCA GCT 1536
 Gly Pro Gln Pro Gly Gln Gln Gln Asn Gly Phe Thr Gly Gln Pro Ala
 500 505 510

ACT TAC CAT CAT AAC AGC ACT ACC ACC TGG ACT GGA AGT AGG ACT GCA 1584
 Thr Tyr His His Asn Ser Thr Thr Trp Thr Gly Ser Arg Thr Ala
 515 520 525

CCA TAC ACA CCT AAT TTG CCT CAC CAC CAA AAC GGC CAT CTT CAG CAC 1632
 Pro Tyr Thr Pro Asn Leu Pro His His Gln Asn Gly His Leu Gln His

530	535	540	
CAC CCG CCT ATG CCG CCC CAT CCC GGA CAT TAC TGG CCT GTT CAC AAT			1680
His Pro Pro Met Pro Pro His Pro Gly His Tyr Trp Pro Val His Asn			
545	550	555	560
GAG CTT GCA TTC CAG CCT CCC ATT TCC AAT CAT CCT GCT CCT GAG TAT			1728
Glu Leu Ala Phe Gln Pro Pro Ile Ser Asn His Pro Ala Pro Glu Tyr			
565	570	575	
TGG TGT TCC ATT GCT TAC TTT GAA ATG GAT GTT CAG GTA GGA GAG ACA			1776
Trp Cys Ser Ile Ala Tyr Phe Glu Met Asp Val Gln Val Gly Glu Thr			
580	585	590	
TTT AAG GTT CCT TCA AGC TGC CCT ATT GTT ACT GTT GAT GGA TAC GTG			1824
Phe Lys Val Pro Ser Ser Cys Pro Ile Val Thr Val Asp Gly Tyr Val			
595	600	605	
GAC CCT TCT GGA GGA GAT CGC TTT TGT TTG GGT CAA CTC TCC AAT GTC			1872
Asp Pro Ser Gly Gly Asp Arg Phe Cys Leu Gly Gln Leu Ser Asn Val			
610	615	620	
CAC AGG ACA GAA GCC ATT GAG AGA GCA AGG TTG CAC ATA GGC AAA GGT			1920
His Arg Thr Glu Ala Ile Glu Arg Ala Arg Leu His Ile Gly Lys Gly			
625	630	635	640
GTG CAG TTG GAA TGT AAA GGT GAA GGT GAT GTT TGG GTC AGG TGC CTT			1968
Val Gln Leu Glu Cys Lys Gly Glu Gly Asp Val Trp Val Arg Cys Leu			
645	650	655	
AGT GAC CAC GCG GTC TTT GTA CAG AGT TAC TAC TTA GAC AGA GAA GCT			2016
Ser Asp His Ala Val Phe Val Gln Ser Tyr Tyr Leu Asp Arg Glu Ala			
660	665	670	
GGG CGT GCA CCT GGA GAT GCT GTT CAT AAG ATC TAC CCA AGT GCA TAT			2064
Gly Arg Ala Pro Gly Asp Ala Val His Lys Ile Tyr Pro Ser Ala Tyr			
675	680	685	
ATA AAG GTC TTT GAT TTG CGT CAG TGT CAT CGA CAG ATG CAG CAG CAG			2112
Ile Lys Val Phe Asp Leu Arg Gln Cys His Arg Gln Met Gln Gln Gln			
690	695	700	
GCG GCT ACT GCA CAA GCT GCA GCA GCT GCC CAG GCA GCA GCC GTG GCA			2160
Ala Ala Thr Ala Gln Ala Ala Ala Ala Ala Gln Ala Ala Val Ala			
705	710	715	720
GGA AAC ATC CCT GGC CCA GGA TCA GTA GGT GGA ATA GCT CCA GCT ATC			2208
Gly Asn Ile Pro Gly Pro Gly Ser Val Gly Gly Ile Ala Pro Ala Ile			
725	730	735	
AGT CTG TCA GCT GCT GCT GGA ATT GGT GTT GAT GAC CTT CGT CGC TTA			2256
Ser Leu Ser Ala Ala Ala Gly Ile Gly Val Asp Asp Leu Arg Arg Leu			
740	745	750	
TGC ATA CTC AGG ATG AGT TTT GTG AAA GGC TGG GGA CCG GAT TAC CCA			2304
Cys Ile Leu Arg Met Ser Phe Val Lys Gly Trp Gly Pro Asp Tyr Pro			
755	760	765	

AGA CAG AGC ATC AAA GAA ACA CCT TGC TGG ATT GAA ATT CAC TTA CAC 2352
 Arg Gln Ser Ile Lys Glu Thr Pro Cys Trp Ile Glu Ile His Leu His
 770 775 780

CGG GCC CTC CAG CTC CTA GAC GAA GTA CTT CAT ACC ATG CCG ATT GCA 2400
 Arg Ala Leu Gln Leu Leu Asp Glu Val Leu His Thr Met Pro Ile Ala
 785 790 795 800

GAC CCA CAA CCT TTA GAC TGA 2421
 Asp Pro Gln Pro Leu Asp
 805

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 806 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10 15
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 20 25 30
 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 35 40 45
 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50 55 60
 Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 65 70 75 80
 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85 90 95
 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 100 105 110
 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 115 120 125
 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 130 135 140
 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
 145 150 155 160
 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
 165 170 175
 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
 180 185 190
 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
 195 200 205
 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
 210 215 220
 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
 225 230 235 240
 Gly Leu Arg Ser Arg Ala Gln Ala Ser Asn Ser Asn Ser Thr Met Asp
 245 250 255

Asn Met Ser Ile Thr Asn Thr Pro Thr Ser Asn Asp Ala Cys Leu Ser	260	265	270
Ile Val His Ser Leu Met Cys His Arg Gln Gly Gly Glu Ser Glu Thr	275	280	285
Phe Ala Lys Arg Ala Ile Glu Ser Leu Val Lys Lys Leu Lys Glu Lys	290	295	300
Lys Asp Glu Leu Asp Ser Leu Ile Thr Ala Ile Thr Thr Asn Gly Ala	305	310	315
His Pro Ser Lys Cys Val Thr Ile Gln Arg Thr Leu Asp Gly Arg Leu	325	330	335
Gln Val Ala Gly Arg Lys Gly Phe Pro His Val Ile Tyr Ala Arg Leu	340	345	350
Trp Arg Trp Pro Asp Leu His Lys Asn Glu Leu Lys His Val Lys Tyr	355	360	365
Cys Gln Tyr Ala Phe Asp Leu Lys Cys Asp Ser Val Cys Val Asn Pro	370	375	380
Tyr His Tyr Glu Arg Val Ser Pro Gly Ile Asp Leu Ser Gly Leu	385	390	395
Thr Leu Gln Ser Asn Ala Pro Ser Ser Met Met Val Lys Asp Glu Tyr	405	410	415
Val His Asp Phe Glu Gly Gln Pro Ser Leu Ser Thr Glu Gly His Ser	420	425	430
Ile Gln Thr Ile Gln His Pro Pro Ser Asn Arg Ala Ser Thr Glu Thr	435	440	445
Tyr Ser Thr Pro Ala Leu Leu Ala Pro Ser Glu Ser Asn Ala Thr Ser	450	455	460
Thr Ala Asn Phe Pro Asn Ile Pro Val Ala Ser Thr Ser Gln Pro Ala	465	470	475
Ser Ile Leu Gly Gly Ser His Ser Glu Gly Leu Leu Gln Ile Ala Ser	485	490	495
Gly Pro Gln Pro Gly Gln Gln Gln Asn Gly Phe Thr Gly Gln Pro Ala	500	505	510
Thr Tyr His His Asn Ser Thr Thr Thr Trp Thr Gly Ser Arg Thr Ala	515	520	525
Pro Tyr Thr Pro Asn Leu Pro His His Gln Asn Gly His Leu Gln His	530	535	540
His Pro Pro Met Pro Pro His Pro Gly His Tyr Trp Pro Val His Asn	545	550	555
Glu Leu Ala Phe Gln Pro Pro Ile Ser Asn His Pro Ala Pro Glu Tyr	565	570	575
Trp Cys Ser Ile Ala Tyr Phe Glu Met Asp Val Gln Val Gly Glu Thr	580	585	590
Phe Lys Val Pro Ser Ser Cys Pro Ile Val Thr Val Asp Gly Tyr Val	595	600	605
Asp Pro Ser Gly Gly Asp Arg Phe Cys Leu Gly Gln Leu Ser Asn Val	610	615	620
His Arg Thr Glu Ala Ile Glu Arg Ala Arg Leu His Ile Gly Lys Gly	625	630	635
Val Gln Leu Glu Cys Lys Gly Glu Gly Asp Val Trp Val Arg Cys Leu	645	650	655
Ser Asp His Ala Val Phe Val Gln Ser Tyr Tyr Leu Asp Arg Glu Ala	660	665	670
Gly Arg Ala Pro Gly Asp Ala Val His Lys Ile Tyr Pro Ser Ala Tyr	675	680	685
Ile Lys Val Phe Asp Leu Arg Gln Cys His Arg Gln Met Gln Gln Gln	690	695	700
Ala Ala Thr Ala Gln Ala Ala Ala Ala Gln Ala Ala Val Ala	705	710	715
			720

ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	48
Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
1				5					10					15		
GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	96
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			20					25					30			
GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	AAG	TTC	ATC	144
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
		35					40					45				
TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	192
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
	50					55					60					
CTG	ACC	TAC	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	CAC	ATG	AAG	240
Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	
65				70					75					80		
CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	288
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	
			85						90				95			
CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	336
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	
			100					105					110			

GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly 115 120 125	384
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr 130 135 140	432
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn 145 150 155 160	480
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser 165 170 175	528
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190	576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205	624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220	672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser 225 230 235 240	720
GGA CTC AGA TCT ACC ATG GCG GGC TGG ATC CAG GCC CAG CAG CTG CAG Gly Leu Arg Ser Thr Met Ala Gly Trp Ile Gln Ala Gln Gln Leu Gln 245 250 255	768
GGA GAC GCG CTG CGC CAG ATG CAG GTG CTG TAC GGC CAG CAC TTC CCC Gly Asp Ala Leu Arg Gln Met Gln Val Leu Tyr Gly Gln His Phe Pro 260 265 270	816
ATC GAG GTC CGG CAC TAC TTG GCC CAG TGG ATT GAG AGC CAG CCA TGG Ile Glu Val Arg His Tyr Leu Ala Gln Trp Ile Glu Ser Gln Pro Trp 275 280 285	864
GAT GCC ATT GAC TTG GAC AAT CCC CAG GAC AGA GCC CAA GCC ACC CAG Asp Ala Ile Asp Leu Asp Asn Pro Gln Asp Arg Ala Gln Ala Thr Gln 290 295 300	912
CTC CTG GAG GGC CTG GTG CAG GAG CTG CAG AAG AAG GCG GAG CAC CAG Leu Leu Glu Gly Leu Val Gln Glu Leu Gln Lys Lys Ala Glu His Gln 305 310 315 320	960
GTG GGG GAA GAT GGG TTT TTA CTG AAG ATC AAG CTG GGG CAC TAC GCC Val Gly Glu Asp Gly Phe Leu Leu Lys Ile Lys Leu Gly His Tyr Ala 325 330 335	1008
ACG CAG CTC CAG AAA ACA TAT GAC CGC TGC CCC CTG GAG CTG GTC CGC Thr Gln Leu Gln Lys Thr Tyr Asp Arg Cys Pro Leu Glu Leu Val Arg	1056

340										345										350										
TGC	ATC	CGG	CAC	ATT	CTG	TAC	AAT	GAA	CAG	AGG	CTG	GTC	CGA	GAA	GCC	1104														
Cys	Ile	Arg	His	Ile	Leu	Tyr	Asn	Glu	Gln	Arg	Leu	Val	Arg	Glu	Ala															
355					360					365																				
AAC	AAT	TGC	AGC	TCT	CCG	GCT	GGG	ATC	CTG	GTT	GAC	GCC	ATG	TCC	CAG	1152														
Asn	Asn	Cys	Ser	Ser	Pro	Ala	Gly	Ile	Leu	Val	Asp	Ala	Met	Ser	Gln															
370					375					380																				
AAG	CAC	CTT	CAG	ATC	AAC	CAG	ACA	TTT	GAG	GAG	CTG	CGA	CTG	GTC	ACG	1200														
Lys	His	Leu	Gln	Ile	Asn	Gln	Thr	Phe	Glu	Glu	Leu	Arg	Leu	Val	Thr															
385					390					395					400															
CAG	GAC	ACA	GAG	AAT	GAG	CTG	AAG	AAA	CTG	CAG	CAG	ACT	CAG	GAG	TAC	1248														
Gln	Asp	Thr	Glu	Asn	Glu	Leu	Lys	Lys	Leu	Gln	Gln	Thr	Gln	Glu	Tyr															
405					410					415																				
TTC	ATC	ATC	CAG	TAC	CAG	GAG	AGC	CTG	AGG	ATC	CAA	GCT	CAG	TTT	GCC	1296														
Phe	Ile	Ile	Gln	Tyr	Gln	Glu	Ser	Leu	Arg	Ile	Gln	Ala	Gln	Phe	Ala															
420					425					430																				
CAG	CTG	GCC	CAG	CTG	AGC	CCC	CAG	GAG	CGT	CTG	AGC	CGG	GAG	ACG	GCC	1344														
Gln	Leu	Ala	Gln	Leu	Ser	Pro	Gln	Glu	Arg	Leu	Ser	Arg	Glu	Thr	Ala															
435					440					445																				
CTC	CAG	CAG	AAG	CAG	GTG	TCT	CTG	GAG	GCC	TGG	TTG	CAG	CGT	GAG	GCA	1392														
Leu	Gln	Gln	Lys	Gln	Val	Ser	Leu	Glu	Ala	Trp	Leu	Gln	Arg	Glu	Ala															
450					455					460																				
CAG	ACA	CTG	CAG	CAG	TAC	CGC	GTG	GAG	CTG	GCC	GAG	AAG	CAC	CAG	AAG	1440														
Gln	Thr	Leu	Gln	Gln	Tyr	Arg	Val	Glu	Leu	Ala	Glu	Lys	His	Gln	Lys															
465					470					475					480															
ACC	CTG	CAG	CTG	CTG	CGG	AAG	CAG	CAG	ACC	ATC	ATC	CTG	GAT	GAC	GAG	1488														
Thr	Leu	Gln	Leu	Leu	Arg	Lys	Gln	Gln	Thr	Ile	Ile	Leu	Asp	Asp	Glu															
485					490					495																				
CTG	ATC	CAG	TGG	AAG	CGG	CGG	CAG	CAG	CTG	GCC	GGG	AAC	GGC	GGG	CCC	1536														
Leu	Ile	Gln	Trp	Lys	Arg	Arg	Gln	Gln	Leu	Ala	Gly	Asn	Gly	Gly	Pro															
500					505					510																				
CCC	GAG	GGC	AGC	CTG	GAC	GTG	CTA	CAG	TCC	TGG	TGT	GAG	AAG	TTG	GCC	1584														
Pro	Glu	Gly	Ser	Leu	Asp	Val	Leu	Gln	Ser	Trp	Cys	Glu	Lys	Leu	Ala															
515					520					525																				
GAG	ATC	ATC	TGG	CAG	AAC	CGG	CAG	CAG	ATC	CGC	AGG	GCT	GAG	CAC	CTC	1632														
Glu	Ile	Ile	Trp	Gln	Asn	Arg	Gln	Gln	Ile	Arg	Arg	Ala	Glu	His	Leu															
530					535					540																				
TGC	CAG	CAG	CTG	CCC	ATC	CCC	GGC	CCA	GTG	GAG	GAG	ATG	CTG	GCC	GAG	1680														
Cys	Gln	Gln	Leu	Pro	Ile	Pro	Gly	Pro	Val	Glu	Glu	Met	Leu	Ala	Glu															
545					550					555					560															
GTC	AAC	GCC	ACC	ATC	ACG	GAC	ATT	ATC	TCA	GCC	CTG	GTG	ACC	AGC	ACA	1728														
Val	Asn	Ala	Thr	Ile	Thr	Asp	Ile	Ile	Ser	Ala	Leu	Val	Thr	Ser	Thr															
565					570					575																				

55

TTC ATC ATT GAG AAG CAG CCT CCT CAG GTC CTG AAG ACC CAG ACC AAG Phe Ile Ile Glu Lys Gln Pro Pro Gln Val Leu Lys Thr Gln Thr Lys 580 585 590	1776
TTT GCA GCC ACC GTA CGC CTG CTG GTG GGC GGG AAG CTG AAC GTG CAC Phe Ala Ala Thr Val Arg Leu Leu Val Gly Gly Lys Leu Asn Val His 595 600 605	1824
ATG AAT CCC CCC CAG GTG AAG GCC ACC ATC ATC AGT GAG CAG CAG GCC Met Asn Pro Pro Gln Val Lys Ala Thr Ile Ile Ser Glu Gln Gln Ala 610 615 620	1872
AAG TCT CTG CTT AAA AAT GAG AAC ACC CGC AAC GAG TGC AGT GGT GAG Lys Ser Leu Leu Lys Asn Glu Asn Thr Arg Asn Glu Cys Ser Gly Glu 625 630 635 640	1920
ATC CTG AAC AAC TGC TGC GTG ATG GAG TAC CAC CAA GCC ACG GGC ACC Ile Leu Asn Asn Cys Cys Val Met Glu Tyr His Gln Ala Thr Gly Thr 645 650 655	1968
CTC AGT GCC CAC TTC AGG AAC ATG TCA CTG AAG AGG ATC AAG CGT GCT Leu Ser Ala His Phe Arg Asn Met Ser Leu Lys Arg Ile Lys Arg Ala 660 665 670	2016
GAC CGG CGG GGT GCA GAG TCC GTG ACA GAG GAG AAG TTC ACA GTC CTG Asp Arg Arg Gly Ala Glu Ser Val Thr Glu Glu Lys Phe Thr Val Leu 675 680 685	2064
TTT GAG TCT CAG TTC AGT GTT GGC AGC AAT GAG CTT GTG TTC CAG GTG Phe Glu Ser Gln Phe Ser Val Gly Ser Asn Glu Leu Val Phe Gln Val 690 695 700	2112
AAG ACT CTG TCC CTA CCT GTG GTT GTC ATC GTC CAC GGC AGC CAG GAC Lys Thr Leu Ser Leu Pro Val Val Val Ile Val His Gly Ser Gln Asp 705 710 715 720	2160
CAC AAT GCC ACG GCT ACT GTG CTG TGG GAC AAT GCC TTT GCT GAG CCG His Asn Ala Thr Ala Thr Val Leu Trp Asp Asn Ala Phe Ala Glu Pro 725 730 735	2208
GGC AGG GTG CCA TTT GCC GTG CCT GAC AAA GTG CTG TGG CCG CAG CTG Gly Arg Val Pro Phe Ala Val Pro Asp Lys Val Leu Trp Pro Gln Leu 740 745 750	2256
TGT GAG GCG CTC AAC ATG AAA TTC AAG GCC GAA GTG CAG AGC AAC CGG Cys Glu Ala Leu Asn Met Lys Phe Lys Ala Glu Val Gln Ser Asn Arg 755 760 765	2304
GGC CTG ACC AAG GAG AAC CTC GTG TTC CTG GCG CAG AAA CTG TTC AAC Gly Leu Thr Lys Glu Asn Leu Val Phe Leu Ala Gln Lys Leu Phe Asn 770 775 780	2352
AAC AGC AGC AGC CAC CTG GAG GAC TAC AGT GGC CTG TCC GTG TCC TGG Asn Ser Ser Ser His Leu Glu Asp Tyr Ser Gly Leu Ser Val Ser Trp 785 790 795 800	2400
TCC CAG TTC AAC AGG GAG AAC TTG CCG GGC TGG AAC TAC ACC TTC TGG Ser Gln Phe Asn Arg Glu Asn Leu Pro Gly Trp Asn Tyr Thr Phe Trp	2448

805	810	815	
CAG TGG TTT GAC GGG GTG ATG GAG GTG TTG AAG AAG CAC CAC AAG CCC Gln Trp Phe Asp Gly Val Met Glu Val Leu Lys Lys His His Lys Pro 820 825 830			2496
CAC TGG AAT GAT GGG GCC ATC CTA GGT TTT GTG AAT AAG CAA CAG GCC His Trp Asn Asp Gly Ala Ile Leu Gly Phe Val Asn Lys Gln Gln Ala 835 840 845			2544
CAC GAC CTG CTC ATC AAC AAG CCC GAC GGG ACC TTC TTG TTG CGC TTT His Asp Leu Leu Ile Asn Lys Pro Asp Gly Thr Phe Leu Leu Arg Phe 850 855 860			2592
AGT GAC TCA GAA ATC GGG GGC ATC ACC ATC GCC TGG AAG TTT GAC TCC Ser Asp Ser Glu Ile Gly Gly Ile Thr Ile Ala Trp Lys Phe Asp Ser 865 870 875 880			2640
CCG GAA CGC AAC CTG TGG AAC CTG AAA CCA TTC ACC ACG CGG GAT TTC Pro Glu Arg Asn Leu Trp Asn Leu Lys Pro Phe Thr Thr Arg Asp Phe 885 890 895			2688
TCC ATC AGG TCC CTG GCT GAC CGG CTG GGG GAC CTG AGC TAT CTC ATC Ser Ile Arg Ser Leu Ala Asp Arg Leu Gly Asp Leu Ser Tyr Leu Ile 900 905 910			2736
TAT GTG TTT CCT GAC CGC CCC AAG GAT GAG GTC TTC TCC AAG TAC TAC Tyr Val Phe Pro Asp Arg Pro Lys Asp Glu Val Phe Ser Lys Tyr Tyr 915 920 925			2784
ACT CCT GTG CTG GCT AAA GCT GTT GAT GGA TAT GTG AAA CCA CAG ATC Thr Pro Val Leu Ala Lys Ala Val Asp Gly Tyr Val Lys Pro Gln Ile 930 935 940			2832
AAG CAA GTG GTC CCT GAG TTT GTG AAT GCA TCT GCA GAT GCT GGG GGC Lys Gln Val Val Pro Glu Phe Val Asn Ala Ser Ala Asp Ala Gly Gly 945 950 955 960			2880
AGC AGC GCC ACG TAC ATG GAC CAG GCC CCC TCC CCA GCT GTG TGC CCC Ser Ser Ala Thr Tyr Met Asp Gln Ala Pro Ser Pro Ala Val Cys Pro 965 970 975			2928
CAG GCT CCC TAT AAC ATG TAC CCA CAG AAC CCT GAC CAT GTA CTC GAT Gln Ala Pro Tyr Asn Met Tyr Pro Gln Asn Pro Asp His Val Leu Asp 980 985 990			2976
CAG GAT GGA GAA TTC GAC CTG GAT GAG ACC ATG GAT GTG GCC AGG CAC Gln Asp Gly Glu Phe Asp Leu Asp Glu Thr Met Asp Val Ala Arg His 995 1000 1005			3024
GTG GAG GAA CTC TTA CGC CGA CCA ATG GAC AGT CTT GAC TCC CGC CTC Val Glu Glu Leu Leu Arg Arg Pro Met Asp Ser Leu Asp Ser Arg Leu 1010 1015 1020			3072
TCG CCC CCT GCC GGT CTT TTC ACC TCT GCC AGA GGC TCC CTC TCA TGA Ser Pro Pro Ala Gly Leu Phe Thr Ser Ala Arg Gly Ser Leu Ser 1025 1030 1035 1			3120

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1039 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1           5           10           15
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
      20           25           30
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
      35           40           45
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
      50           55           60
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
      65           70           75           80
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
      85           90           95
Arg Thr Ile Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
      100          105          110
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
      115          120          125
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
      130          135          140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
      145          150          155          160
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
      165          170          175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
      180          185          190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
      195          200          205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
      210          215          220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
      225          230          235          240
Gly Leu Arg Ser Thr Met Ala Gly Trp Ile Gln Ala Gln Gln Leu Gln
      245          250          255
Gly Asp Ala Leu Arg Gln Met Gln Val Leu Tyr Gly Gln His Phe Pro
      260          265          270
Ile Glu Val Arg His Tyr Leu Ala Gln Trp Ile Glu Ser Gln Pro Trp
      275          280          285
Asp Ala Ile Asp Leu Asp Asn Pro Gln Asp Arg Ala Gln Ala Thr Gln
      290          295          300
Leu Leu Glu Gly Leu Val Gln Glu Leu Gln Lys Lys Ala Glu His Gln
      305          310          315          320
Val Gly Glu Asp Gly Phe Leu Leu Lys Ile Lys Leu Gly His Tyr Ala
      325          330          335
Thr Gln Leu Gln Lys Thr Tyr Asp Arg Cys Pro Leu Glu Leu Val Arg
      340          345          350

```


Cys	Ile	Arg	His	Ile	Leu	Tyr	Asn	Glu	Gln	Arg	Leu	Val	Arg	Glu	Ala	355	360	365
Asn	Asn	Cys	Ser	Ser	Pro	Ala	Gly	Ile	Leu	Val	Asp	Ala	Met	Ser	Gln	370	375	380
Lys	His	Leu	Gln	Ile	Asn	Gln	Thr	Phe	Glu	Glu	Leu	Arg	Leu	Val	Thr	385	390	395
Gln	Asp	Thr	Glu	Asn	Glu	Leu	Lys	Lys	Leu	Gln	Gln	Thr	Gln	Glu	Tyr	405	410	415
Phe	Ile	Ile	Gln	Tyr	Gln	Glu	Ser	Leu	Arg	Ile	Gln	Ala	Gln	Phe	Ala	420	425	430
Gln	Leu	Ala	Gln	Leu	Ser	Pro	Gln	Glu	Arg	Leu	Ser	Arg	Glu	Thr	Ala	435	440	445
Leu	Gln	Gln	Lys	Gln	Val	Ser	Leu	Glu	Ala	Trp	Leu	Gln	Arg	Glu	Ala	450	455	460
Gln	Thr	Leu	Gln	Gln	Tyr	Arg	Val	Glu	Leu	Ala	Glu	Lys	His	Gln	Lys	465	470	475
Thr	Leu	Gln	Leu	Leu	Arg	Lys	Gln	Gln	Thr	Ile	Ile	Leu	Asp	Asp	Glu	485	490	495
Leu	Ile	Gln	Trp	Lys	Arg	Arg	Gln	Gln	Leu	Ala	Gly	Asn	Gly	Gly	Pro	500	505	510
Pro	Glu	Gly	Ser	Leu	Asp	Val	Leu	Gln	Ser	Trp	Cys	Glu	Lys	Leu	Ala	515	520	525
Glu	Ile	Ile	Trp	Gln	Asn	Arg	Gln	Gln	Ile	Arg	Arg	Ala	Glu	His	Leu	530	535	540
Cys	Gln	Gln	Leu	Pro	Ile	Pro	Gly	Pro	Val	Glu	Glu	Met	Leu	Ala	Glu	545	550	555
Val	Asn	Ala	Thr	Ile	Thr	Asp	Ile	Ile	Ser	Ala	Leu	Val	Thr	Ser	Thr	565	570	575
Phe	Ile	Ile	Glu	Lys	Gln	Pro	Pro	Gln	Val	Leu	Lys	Thr	Gln	Thr	Lys	580	585	590
Phe	Ala	Ala	Thr	Val	Arg	Leu	Leu	Val	Gly	Gly	Lys	Leu	Asn	Val	His	595	600	605
Met	Asn	Pro	Pro	Gln	Val	Lys	Ala	Thr	Ile	Ile	Ser	Glu	Gln	Gln	Ala	610	615	620
Lys	Ser	Leu	Leu	Lys	Asn	Glu	Asn	Thr	Arg	Asn	Glu	Cys	Ser	Gly	Glu	625	630	635
Ile	Leu	Asn	Asn	Cys	Cys	Val	Met	Glu	Tyr	His	Gln	Ala	Thr	Gly	Thr	645	650	655
Leu	Ser	Ala	His	Phe	Arg	Asn	Met	Ser	Leu	Lys	Arg	Ile	Lys	Arg	Ala	660	665	670
Asp	Arg	Arg	Gly	Ala	Glu	Ser	Val	Thr	Glu	Glu	Lys	Phe	Thr	Val	Leu	675	680	685
Phe	Glu	Ser	Gln	Phe	Ser	Val	Gly	Ser	Asn	Glu	Leu	Val	Phe	Gln	Val	690	695	700
Lys	Thr	Leu	Ser	Leu	Pro	Val	Val	Val	Ile	Val	His	Gly	Ser	Gln	Asp	705	710	715
His	Asn	Ala	Thr	Ala	Thr	Val	Leu	Trp	Asp	Asn	Ala	Phe	Ala	Glu	Pro	725	730	735
Gly	Arg	Val	Pro	Phe	Ala	Val	Pro	Asp	Lys	Val	Leu	Trp	Pro	Gln	Leu	740	745	750
Cys	Glu	Ala	Leu	Asn	Met	Lys	Phe	Lys	Ala	Glu	Val	Gln	Ser	Asn	Arg	755	760	765
Gly	Leu	Thr	Lys	Glu	Asn	Leu	Val	Phe	Leu	Ala	Gln	Lys	Leu	Phe	Asn	770	775	780
Asn	Ser	Ser	Ser	His	Leu	Glu	Asp	Tyr	Ser	Gly	Leu	Ser	Val	Ser	Trp	785	790	795
Ser	Gln	Phe	Asn	Arg	Glu	Asn	Leu	Pro	Gly	Trp	Asn	Tyr	Thr	Phe	Trp	805	810	815

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1875 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...1872
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

ATG	GCG	GCG	GCG	GCG	GCG	GCT	CCG	GGG	GGC	GGG	GGC	GGG	GAG	CCC	AGG	48
Met	Ala	Ala	Ala	Ala	Ala	Ala	Pro	Gly	Gly	Gly	Gly	Gly	Glu	Pro	Arg	
1				5					10					15		
GGA	ACT	GCT	GGG	GTC	GTC	CCG	GTG	GTC	CCC	GGG	GAG	GTG	GAG	GTG	GTG	96
Gly	Thr	Ala	Gly	Val	Val	Pro	Val	Val	Pro	Gly	Glu	Val	Glu	Val	Val	
			20					25					30			
AAG	GGG	CAG	CCA	TTC	GAT	GTG	GGC	CCA	CGC	TAC	ACG	CAG	CTG	CAG	TAC	144
Lys	Gly	Gln	Pro	Phe	Asp	Val	Gly	Pro	Arg	Tyr	Thr	Gln	Leu	Gln	Tyr	
	35						40					45				

ATC	GGC	GAG	GGC	GCG	TAC	GGC	ATG	GTC	AGC	TCA	GCT	TAT	GAC	CAC	GTG	192
Ile	Gly	Glu	Gly	Ala	Tyr	Gly	Met	Val	Ser	Ser	Ala	Tyr	Asp	His	Val	
50						55					60					
CGC	AAG	ACC	AGA	GTG	GCC	ATC	AAG	AAG	ATC	AGC	CCC	TTT	GAG	CAT	CAA	240
Arg	Lys	Thr	Arg	Val	Ala	Ile	Lys	Lys	Ile	Ser	Pro	Phe	Glu	His	Gln	
65					70					75					80	
ACC	TAC	TGT	CAG	CGC	ACG	CTG	AGG	GAG	ATC	CAG	ATC	TTG	CTG	CGA	TTC	288
Thr	Tyr	Cys	Gln	Arg	Thr	Leu	Arg	Glu	Ile	Gln	Ile	Leu	Leu	Arg	Phe	
			85						90						95	
CGC	CAT	GAG	AAT	GTT	ATA	GGC	ATC	CGA	GAC	ATC	CTC	AGA	GCG	CCC	ACC	336
Arg	His	Glu	Asn	Val	Ile	Gly	Ile	Arg	Asp	Ile	Leu	Arg	Ala	Pro	Thr	
			100					105					110			
CTG	GAA	GCC	ATG	AGA	GAT	GTT	TAC	ATT	GTT	CAG	GAC	CTC	ATG	GAG	ACA	384
Leu	Glu	Ala	Met	Arg	Asp	Val	Tyr	Ile	Val	Gln	Asp	Leu	Met	Glu	Thr	
			115				120					125				
GAC	CTG	TAC	AAG	CTG	CTT	AAA	AGC	CAG	CAG	CTG	AGC	AAT	GAC	CAC	ATC	432
Asp	Leu	Tyr	Lys	Leu	Leu	Lys	Ser	Gln	Gln	Leu	Ser	Asn	Asp	His	Ile	
	130					135						140				
TGC	TAC	TTC	CTC	TAC	CAG	ATC	CTC	CGG	GGC	CTC	AAG	TAT	ATA	CAC	TCA	480
Cys	Tyr	Phe	Leu	Tyr	Gln	Ile	Leu	Arg	Gly	Leu	Lys	Tyr	Ile	His	Ser	
145					150					155					160	
GCC	AAT	GTG	CTG	CAC	CGG	GAC	CTG	AAG	CCT	TCC	AAT	CTG	CTT	ATC	AAC	528
Ala	Asn	Val	Leu	His	Arg	Asp	Leu	Lys	Pro	Ser	Asn	Leu	Leu	Ile	Asn	
				165					170					175		
ACC	ACC	TGC	GAC	CTT	AAG	ATC	TGT	GAT	TTT	GGC	CTG	GCC	CGG	ATT	GCT	576
Thr	Thr	Cys	Asp	Leu	Lys	Ile	Cys	Asp	Phe	Gly	Leu	Ala	Arg	Ile	Ala	
			180					185					190			
GAC	CCT	GAG	CAC	GAC	CAC	ACT	GGC	TTT	CTG	ACG	GAG	TAT	GTG	GCC	ACA	624
Asp	Pro	Glu	His	Asp	His	Thr	Gly	Phe	Leu	Thr	Glu	Tyr	Val	Ala	Thr	
		195					200					205				
CGC	TGG	TAC	CGA	GCC	CCA	GAG	ATC	ATG	CTT	AAT	TCC	AAG	GGC	TAC	ACC	672
Arg	Trp	Tyr	Arg	Ala	Pro	Glu	Ile	Met	Leu	Asn	Ser	Lys	Gly	Tyr	Thr	
	210					215					220					
AAA	TCC	ATC	GAC	ATC	TGG	TCT	GTG	GGC	TGC	ATT	CTG	GCT	GAG	ATG	CTC	720
Lys	Ser	Ile	Asp	Ile	Trp	Ser	Val	Gly	Cys	Ile	Leu	Ala	Glu	Met	Leu	
225					230					235					240	
TCC	AAC	CGG	CCC	ATC	TTC	CCC	GGC	AAG	CAC	TAC	CTG	GAC	CAG	CTC	AAC	768
Ser	Asn	Arg	Pro	Ile	Phe	Pro	Gly	Lys	His	Tyr	Leu	Asp	Gln	Leu	Asn	
				245					250					255		
CAC	ATT	CTA	GGT	ATC	TTG	GGT	TCC	CCA	TCC	CAG	GAG	GAC	CTT	AAT	TGC	816
His	Ile	Leu	Gly	Ile	Leu	Gly	Ser	Pro	Ser	Gln	Glu	Asp	Leu	Asn	Cys	
			260					265					270			
ATC	ATT	AAC	ATG	AAG	GCC	CGA	AAC	TAC	CTG	CAG	TCT	CTG	CCC	TCG	AAA	864
Ile	Ile	Asn	Met	Lys	Ala	Arg	Asn	Tyr	Leu	Gln	Ser	Leu	Pro	Ser	Lys	

275	280	285	
ACC AAG GTG GCT TGG GCC AAG CTC TTT CCT AAA TCT GAC TCC AAA GCT Thr Lys Val Ala Trp Ala Lys Leu Phe Pro Lys Ser Asp Ser Lys Ala 290 295 300			912
CTT GAC CTG CTG GAC CGG ATG TTA ACC TTC AAC CCA AAC AAG CGC ATC Leu Asp Leu Leu Asp Arg Met Leu Thr Phe Asn Pro Asn Lys Arg Ile 305 310 315 320			960
ACA GTA GAG GAA GCG CTG GCT CAC CCT TAC CTG GAA CAG TAC TAC GAT Thr Val Glu Glu Ala Leu Ala His Pro Tyr Leu Glu Gln Tyr Tyr Asp 325 330 335			1008
CCG ACA GAT GAG CCA GTG GCC GAG GAG CCA TTC ACC TTC GAC ATG GAG Pro Thr Asp Glu Pro Val Ala Glu Glu Pro Phe Thr Phe Asp Met Glu 340 345 350			1056
CTG GAT GAC CTC CCC AAG GAG CGG CTG AAG GAG TTG ATC TTC CAG GAG Leu Asp Asp Leu Pro Lys Glu Arg Leu Lys Glu Leu Ile Phe Gln Glu 355 360 365			1104
ACA GCC CGC TTC CAG CCA GGG GCG CCA GAG GGC CCC GGG CGC GCC ATG Thr Ala Arg Phe Gln Pro Gly Ala Pro Glu Gly Pro Gly Arg Ala Met 370 375 380			1152
AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT GTC CCA ATT CTT GTT GAA Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu 385 390 395 400			1200
TTA GAT GGC GAT GTT AAT GGG CAA AAA TTC TCT GTT AGT GGA GAG GGT Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu Gly 405 410 415			1248
GAA GGT GAT GCA ACA TAC GGA AAA CTT ACC CTT AAA TTT ATT TGC ACT Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr 420 425 430			1296
ACT GGG AAG CTA CCT GTT CCA TGG CCA ACG CTT GTC ACT ACT CTC ACT Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr 435 440 445			1344
TAT GGT GTT CAA TGC TTT TCT AGA TAC CCA GAT CAT ATG AAA CAG CAT Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His 450 455 460			1392
GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT TAT GTA CAG GAA AGA ACT Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr 465 470 475 480			1440
ATA TTT TAC AAA GAT GAC GGG AAC TAC AAG ACA CGT GCT GAA GTC AAG Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys 485 490 495			1488
TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG TTA AAA GGT ATT GAT Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp 500 505 510			1536

TTT AAA GAA GAT GGA AAC ATT CTT GGA CAC AAA ATG GAA TAC AAT TAT 1584
 Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn Tyr
 515 520 525
 AAC TCA CAT AAT GTA TAC ATC ATG GCA GAC AAA CCA AAG AAT GGC ATC 1632
 Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro Lys Asn Gly Ile
 530 535 540
 AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT AAA GAT GGA AGC GTT CAA 1680
 Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val Gln
 545 550 555 560
 TTA GCA GAC CAT TAT CAA CAA AAT ACT CCA ATT GGC GAT GGC CCT GTC 1728
 Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val
 565 570 575
 CTT TTA CCA GAC AAC CAT TAC CTG TCC ACG CAA TCT GCC CTT TCC AAA 1776
 Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys
 580 585 590
 GAT CCC AAC GAA AAG AGA GAT CAC ATG ATC CTT CTT GAG TTT GTA ACA 1824
 Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val Thr
 595 600 605
 GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA TAC AAA CCT CAG GAG T 1873
 Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys Pro Gln Glu
 610 615 620
 AA 1875

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 624 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Met Ala Ala Ala Ala Ala Pro Gly Gly Gly Gly Glu Pro Arg
 1 5 10 15
 Gly Thr Ala Gly Val Val Pro Val Val Pro Gly Glu Val Glu Val Val
 20 25 30
 Lys Gly Gln Pro Phe Asp Val Gly Pro Arg Tyr Thr Gln Leu Gln Tyr
 35 40 45
 Ile Gly Glu Gly Ala Tyr Gly Met Val Ser Ser Ala Tyr Asp His Val
 50 55 60
 Arg Lys Thr Arg Val Ala Ile Lys Lys Ile Ser Pro Phe Glu His Gln
 65 70 75 80
 Thr Tyr Cys Gln Arg Thr Leu Arg Glu Ile Gln Ile Leu Leu Arg Phe
 85 90 95
 Arg His Glu Asn Val Ile Gly Ile Arg Asp Ile Leu Arg Ala Pro Thr
 100 105 110
 Leu Glu Ala Met Arg Asp Val Tyr Ile Val Gln Asp Leu Met Glu Thr

115	120	125
Asp Leu Tyr Lys Leu Leu Lys Ser Gln Gln Leu Ser Asn Asp His Ile		
130	135	140
Cys Tyr Phe Leu Tyr Gln Ile Leu Arg Gly Leu Lys Tyr Ile His Ser		
145	150	155
Ala Asn Val Leu His Arg Asp Leu Lys Pro Ser Asn Leu Leu Ile Asn		160
165	170	175
Thr Thr Cys Asp Leu Lys Ile Cys Asp Phe Gly Leu Ala Arg Ile Ala		
180	185	190
Asp Pro Glu His Asp His Thr Gly Phe Leu Thr Glu Tyr Val Ala Thr		
195	200	205
Arg Trp Tyr Arg Ala Pro Glu Ile Met Leu Asn Ser Lys Gly Tyr Thr		
210	215	220
Lys Ser Ile Asp Ile Trp Ser Val Gly Cys Ile Leu Ala Glu Met Leu		
225	230	235
Ser Asn Arg Pro Ile Phe Pro Gly Lys His Tyr Leu Asp Gln Leu Asn		240
245	250	255
His Ile Leu Gly Ile Leu Gly Ser Pro Ser Gln Glu Asp Leu Asn Cys		
260	265	270
Ile Ile Asn Met Lys Ala Arg Asn Tyr Leu Gln Ser Leu Pro Ser Lys		
275	280	285
Thr Lys Val Ala Trp Ala Lys Leu Phe Pro Lys Ser Asp Ser Lys Ala		
290	295	300
Leu Asp Leu Leu Asp Arg Met Leu Thr Phe Asn Pro Asn Lys Arg Ile		
305	310	315
Thr Val Glu Glu Ala Leu Ala His Pro Tyr Leu Glu Gln Tyr Tyr Asp		
325	330	335
Pro Thr Asp Glu Pro Val Ala Glu Glu Pro Phe Thr Phe Asp Met Glu		
340	345	350
Leu Asp Asp Leu Pro Lys Glu Arg Leu Lys Glu Leu Ile Phe Gln Glu		
355	360	365
Thr Ala Arg Phe Gln Pro Gly Ala Pro Glu Gly Pro Gly Arg Ala Met		
370	375	380
Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu		
385	390	395
Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu Gly		
405	410	415
Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr		
420	425	430
Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr		
435	440	445
Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His		
450	455	460
Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr		
465	470	475
Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys		
485	490	495
Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp		
500	505	510
Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn Tyr		
515	520	525
Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro Lys Asn Gly Ile		
530	535	540
Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val Gln		
545	550	555
Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val		
565	570	575
Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys		

CAC CGT GAC CTC AAG CCT TCC AAC CTC CTG CTG AAC ACC ACT TGT GAT His Arg Asp Leu Lys Pro Ser Asn Leu Leu Leu Asn Thr Thr Cys Asp 145 150 155 160	480
CTC AAG ATC TGT GAC TTT GGC CTT GCC CGT GTT GCA GAT CCA GAC CAT Leu Lys Ile Cys Asp Phe Gly Leu Ala Arg Val Ala Asp Pro Asp His 165 170 175	528
GAT CAT ACA GGG TTC TTG ACA GAG TAT GTA GCC ACG CGT TGG TAC AGA Asp His Thr Gly Phe Leu Thr Glu Tyr Val Ala Thr Arg Trp Tyr Arg 180 185 190	576
GCT CCA GAA ATT ATG TTG AAT TCC AAG GGT TAT ACC AAG TCC ATT GAT Ala Pro Glu Ile Met Leu Asn Ser Lys Gly Tyr Thr Lys Ser Ile Asp 195 200 205	624
ATT TGG TCT GTG GGC TGC ATC CTG GCA GAG ATG CTA TCC AAC AGG CCT Ile Trp Ser Val Gly Cys Ile Leu Ala Glu Met Leu Ser Asn Arg Pro 210 215 220	672
ATC TTC CCA GGA AAG CAT TAC CTT GAC CAG CTG AAT CAC ATC CTG GGT Ile Phe Pro Gly Lys His Tyr Leu Asp Gln Leu Asn His Ile Leu Gly 225 230 235 240	720
ATT CTT GGA TCT CCA TCA CAG GAA GAT CTG AAT TGT ATA ATA AAT TTA Ile Leu Gly Ser Pro Ser Gln Glu Asp Leu Asn Cys Ile Ile Asn Leu 245 250 255	768
AAA GCT AGA AAC TAT TTG CTT TCT CTC CCG CAC AAA AAT AAG GTG CCG Lys Ala Arg Asn Tyr Leu Leu Ser Leu Pro His Lys Asn Lys Val Pro 260 265 270	816
TGG AAC AGG TTG TTC CCA AAC GCT GAC TCC AAA GCT CTG GAT TTA CTG Trp Asn Arg Leu Phe Pro Asn Ala Asp Ser Lys Ala Leu Asp Leu Leu 275 280 285	864
GAT AAA ATG TTG ACA TTT AAC CCT CAC AAG AGG ATT GAA GTT GAA CAG Asp Lys Met Leu Thr Phe Asn Pro His Lys Arg Ile Glu Val Glu Gln 290 295 300	912
GCT CTG GCC CAC CCG TAC CTG GAG CAG TAT TAT GAC CCA AGT GAT GAG Ala Leu Ala His Pro Tyr Leu Glu Gln Tyr Tyr Asp Pro Ser Asp Glu 305 310 315 320	960
CCC ATT GCT GAA GCA CCA TTC AAG TTT GAC ATG GAG CTG GAC GAC TTA Pro Ile Ala Glu Ala Pro Phe Lys Phe Asp Met Glu Leu Asp Asp Leu 325 330 335	1008
CCT AAG GAG AAG CTC AAA GAA CTC ATT TTT GAA GAG ACT GCT CGA TTC Pro Lys Glu Lys Leu Lys Glu Leu Ile Phe Glu Glu Thr Ala Arg Phe 340 345 350	1056
CAG CCA GGA TAC AGA TCT ATG GAT CCA CCG GTC GCC ACC ATG GTG AGC Gln Pro Gly Tyr Arg Ser Met Asp Pro Pro Val Ala Thr Met Val Ser 355 360 365	1104
AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG	1152

Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu	
370 375 380	
GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG	1200
Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu	
385 390 395 400	
GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC	1248
Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr	
405 410 415	
GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC	1296
Gly Lys Leu Pro Val Pro Trp Thr Leu Val Thr Thr Leu Thr Tyr	
420 425 430	
GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC	1344
Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp	
435 440 445	
TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC	1392
Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile	
450 455 460	
TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC	1440
Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe	
465 470 475 480	
GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC	1488
Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe	
485 490 495	
AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC	1536
Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn	
500 505 510	
AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG	1584
Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys	
515 520 525	
GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC	1632
Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu	
530 535 540	
GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG	1680
Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu	
545 550 555 560	
CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC	1728
Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp	
565 570 575	
CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC	1776
Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala	
580 585 590	
GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AA GTAA	1815
Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys	
595 600	

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 604 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```

Met Ala Ala Ala Ala Ala Gly Pro Glu Met Val Arg Gly Gln Val
 1           5           10           15
Phe Asp Val Gly Pro Arg Tyr Thr Asn Leu Ser Tyr Ile Gly Glu Gly
 20           25           30
Ala Tyr Gly Met Val Cys Ser Ala Tyr Asp Asn Leu Asn Lys Val Arg
 35           40           45
Val Ala Ile Lys Lys Ile Ser Pro Phe Glu His Gln Thr Tyr Cys Gln
 50           55           60
Arg Thr Leu Arg Glu Ile Lys Ile Leu Leu Arg Phe Arg His Glu Asn
 65           70           75           80
Ile Ile Gly Ile Asn Asp Ile Ile Arg Ala Pro Thr Ile Glu Gln Met
 85           90           95
Lys Asp Val Tyr Ile Val Gln Asp Leu Met Glu Thr Asp Leu Tyr Lys
 100          105          110
Leu Leu Lys Thr Gln His Leu Ser Asn Asp His Ile Cys Tyr Phe Leu
 115          120          125
Tyr Gln Ile Leu Arg Gly Leu Lys Tyr Ile His Ser Ala Asn Val Leu
 130          135          140
His Arg Asp Leu Lys Pro Ser Asn Leu Leu Leu Asn Thr Thr Cys Asp
 145          150          155          160
Leu Lys Ile Cys Asp Phe Gly Leu Ala Arg Val Ala Asp Pro Asp His
 165          170          175
Asp His Thr Gly Phe Leu Thr Glu Tyr Val Ala Thr Arg Trp Tyr Arg
 180          185          190
Ala Pro Glu Ile Met Leu Asn Ser Lys Gly Tyr Thr Lys Ser Ile Asp
 195          200          205
Ile Trp Ser Val Gly Cys Ile Leu Ala Glu Met Leu Ser Asn Arg Pro
 210          215          220
Ile Phe Pro Gly Lys His Tyr Leu Asp Gln Leu Asn His Ile Leu Gly
 225          230          235          240
Ile Leu Gly Ser Pro Ser Gln Glu Asp Leu Asn Cys Ile Ile Asn Leu
 245          250          255
Lys Ala Arg Asn Tyr Leu Leu Ser Leu Pro His Lys Asn Lys Val Pro
 260          265          270
Trp Asn Arg Leu Phe Pro Asn Ala Asp Ser Lys Ala Leu Asp Leu Leu
 275          280          285
Asp Lys Met Leu Thr Phe Asn Pro His Lys Arg Ile Glu Val Glu Gln
 290          295          300
Ala Leu Ala His Pro Tyr Leu Glu Gln Tyr Tyr Asp Pro Ser Asp Glu
 305          310          315          320
Pro Ile Ala Glu Ala Pro Phe Lys Phe Asp Met Glu Leu Asp Asp Leu
 325          330          335
Pro Lys Glu Lys Leu Lys Glu Leu Ile Phe Glu Glu Thr Ala Arg Phe

```


ATC CTG AAG TTC CCT CAC ATT AGC CAG TGT GAA GAC CTC CGA AGG ACC Ile Leu Lys Phe Pro His Ile Ser Gln Cys Glu Asp Leu Arg Arg Thr 35 40 45	144
ATA GAC AGA GAT TAC TGC AGT TTA TGT GAC AAG CAG CCA ATC GGG AGG Ile Asp Arg Asp Tyr Cys Ser Leu Cys Asp Lys Gln Pro Ile Gly Arg 50 55 60	192
CTG CTT TTC CGG CAG TTT TGT GAA ACC AGG CCT GGG CTG GAG TGT TAC Leu Leu Phe Arg Gln Phe Cys Glu Thr Arg Pro Gly Leu Glu Cys Tyr 65 70 75 80	240
ATT CAG TTC CTG GAC TCC GTG GCA GAA TAT GAA GTT ACT CCA GAT GAA Ile Gln Phe Leu Asp Ser Val Ala Glu Tyr Glu Val Thr Pro Asp Glu 85 90 95	288
AAA CTG GGA GAG AAA GGG AAG GAA ATT ATG ACC AAG TAC CTC ACC CCA Lys Leu Gly Glu Lys Gly Lys Glu Ile Met Thr Lys Tyr Leu Thr Pro 100 105 110	336
AAG TCC CCT GTT TTC ATA GCC CAA GTT GGC CAA GAC CTG GTC TCC CAG Lys Ser Pro Val Phe Ile Ala Gln Val Gly Gln Asp Leu Val Ser Gln 115 120 125	384
ACG GAG GAG AAG CTC CTA CAG AAG CCG TGC AAA GAA CTC TTT TCT GCC Thr Glu Glu Lys Leu Leu Gln Lys Pro Cys Lys Glu Leu Phe Ser Ala 130 135 140	432
TGT GCA CAG TCT GTC CAC GAG TAC CTG AGG GGA GAA CCA TTC CAC GAA Cys Ala Gln Ser Val His Glu Tyr Leu Arg Gly Glu Pro Phe His Glu 145 150 155 160	480
TAT CTG GAC AGC ATG TTT TTT GAC CGC TTT CTC CAG TGG AAG TGG TTG Tyr Leu Asp Ser Met Phe Phe Asp Arg Phe Leu Gln Trp Lys Trp Leu 165 170 175	528
GAA AGG CAA CCG GTG ACC AAA AAC ACT TTC AGG CAG TAT CGA GTG CTA Glu Arg Gln Pro Val Thr Lys Asn Thr Phe Arg Gln Tyr Arg Val Leu 180 185 190	576
GGA AAA GGG GGC TTC GGG GAG GTC TGT GCC TGC CAG GTT CGG GCC ACG Gly Lys Gly Gly Phe Gly Glu Val Cys Ala Cys Gln Val Arg Ala Thr 195 200 205	624
GGT AAA ATG TAT GCC TGC AAG CGC TTG GAG AAG AAG AGG ATC AAA AAG Gly Lys Met Tyr Ala Cys Lys Arg Leu Glu Lys Lys Arg Ile Lys Lys 210 215 220	672
AGG AAA GGG GAG TCC ATG GCC CTC AAT GAG AAG CAG ATC CTC GAG AAG Arg Lys Gly Glu Ser Met Ala Leu Asn Glu Lys Gln Ile Leu Glu Lys 225 230 235 240	720
GTC AAC AGT CAG TTT GTG GTC AAC CTG GCC TAT GCC TAC GAG ACC AAG Val Asn Ser Gln Phe Val Val Asn Leu Ala Tyr Ala Tyr Glu Thr Lys 245 250 255	768
GAT GCA CTG TGC TTG GTC CTG ACC ATC ATG AAT GGG GGT GAC CTG AAG	816

Asp	Ala	Leu	Cys	Leu	Val	Leu	Thr	Ile	Met	Asn	Gly	Gly	Asp	Leu	Lys	
			260						265						270	
TTC	CAC	ATC	TAC	AAC	ATG	GGC	AAC	CCT	GGC	TTC	GAG	GAG	GAG	CGG	GCC	864
Phe	His	Ile	Tyr	Asn	Met	Gly	Asn	Pro	Gly	Phe	Glu	Glu	Glu	Arg	Ala	
			275					280					285			
TTG	TTT	TAT	GCG	GCA	GAG	ATC	CTC	TGC	GGC	TTA	GAA	GAC	CTC	CAC	CGT	912
Leu	Phe	Tyr	Ala	Ala	Glu	Ile	Leu	Cys	Gly	Leu	Glu	Asp	Leu	His	Arg	
			290				295				300					
GAG	AAC	ACC	GTC	TAC	CGA	GAT	CTG	AAA	CCT	GAA	AAC	ATC	CTG	TTA	GAT	960
Glu	Asn	Thr	Val	Tyr	Arg	Asp	Leu	Lys	Pro	Glu	Asn	Ile	Leu	Leu	Asp	
305					310					315					320	
GAT	TAT	GGC	CAC	ATT	AGG	ATC	TCA	GAC	CTG	GGC	TTG	GCT	GTG	AAG	ATC	1008
Asp	Tyr	Gly	His	Ile	Arg	Ile	Ser	Asp	Leu	Gly	Leu	Ala	Val	Lys	Ile	
				325					330					335		
CCC	GAG	GGA	GAC	CTG	ATC	CGC	GGC	CGG	GTG	GGC	ACT	GTT	GGC	TAC	ATG	1056
Pro	Glu	Gly	Asp	Leu	Ile	Arg	Gly	Arg	Val	Gly	Thr	Val	Gly	Tyr	Met	
			340					345					350			
GCC	CCC	GAA	GTC	CTG	AAC	AAC	CAG	AGG	TAC	GGC	CTG	AGC	CCC	GAC	TAC	1104
Ala	Pro	Glu	Val	Leu	Asn	Asn	Gln	Arg	Tyr	Gly	Leu	Ser	Pro	Asp	Tyr	
			355				360					365				
TGG	GGC	CTT	GGC	TGC	CTC	ATC	TAT	GAG	ATG	ATC	GAG	GGC	CAG	TCG	CCG	1152
Trp	Gly	Leu	Gly	Cys	Leu	Ile	Tyr	Glu	Met	Ile	Glu	Gly	Gln	Ser	Pro	
			370				375				380					
TTC	CGC	GGC	CGT	AAG	GAG	AAG	GTG	AAG	CGG	GAG	GAG	GTG	GAC	CGC	CGG	1200
Phe	Arg	Gly	Arg	Lys	Glu	Lys	Val	Lys	Arg	Glu	Glu	Val	Asp	Arg	Arg	
385					390					395					400	
GTC	CTG	GAG	ACG	GAG	GAG	GTG	TAC	TCC	CAC	AAG	TTC	TCC	GAG	GAG	GCC	1248
Val	Leu	Glu	Thr	Glu	Glu	Val	Tyr	Ser	His	Lys	Phe	Ser	Glu	Glu	Ala	
				405					410					415		
AAG	TCC	ATC	TGC	AAG	ATG	CTG	CTC	ACG	AAA	GAT	GCG	AAG	CAG	AGG	CTG	1296
Lys	Ser	Ile	Cys	Lys	Met	Leu	Leu	Thr	Lys	Asp	Ala	Lys	Gln	Arg	Leu	
			420					425					430			
GGC	TGC	CAG	GAG	GAG	GGG	GCT	GCA	GAG	GTC	AAG	AGA	CAC	CCC	TTC	TTC	1344
Gly	Cys	Gln	Glu	Glu	Gly	Ala	Ala	Glu	Val	Lys	Arg	His	Pro	Phe	Phe	
			435				440					445				
AGG	AAC	ATG	AAC	TTC	AAG	CGC	TTA	GAA	GCC	GGG	ATG	TTG	GAC	CCT	CCC	1392
Arg	Asn	Met	Asn	Phe	Lys	Arg	Leu	Glu	Ala	Gly	Met	Leu	Asp	Pro	Pro	
			450			455					460					
TTC	GTT	CCA	GAC	CCC	CGC	GCT	GTG	TAC	TGT	AAG	GAC	GTG	CTG	GAC	ATC	1440
Phe	Val	Pro	Asp	Pro	Arg	Ala	Val	Tyr	Cys	Lys	Asp	Val	Leu	Asp	Ile	
465					470					475				480		
GAG	CAG	TTC	TCC	ACT	GTG	AAG	GGC	GTC	AAT	CTG	GAC	CAC	ACA	GAC	GAC	1488
Glu	Gln	Phe	Ser	Thr	Val	Lys	Gly	Val	Asn	Leu	Asp	His	Thr	Asp	Asp	
				485				490						495		

GAC TTC TAC TCC AAG TTC TCC ACG GGC TCT GTG TCC ATC CCA TGG CAA	1536
Asp Phe Tyr Ser Lys Phe Ser Thr Gly Ser Val Ser Ile Pro Trp Gln	
500 505 510	
AAC GAG ATG ATA GAA ACA GAA TGC TTT AAG GAG CTG AAC GTG TTT GGA	1584
Asn Glu Met Ile Glu Thr Glu Cys Phe Lys Glu Leu Asn Val Phe Gly	
515 520 525	
CCT AAT GGT ACC CTC CCG CCA GAT CTG AAC AGA AAC CAC CCT CCG GAA	1632
Pro Asn Gly Thr Leu Pro Pro Asp Leu Asn Arg Asn His Pro Pro Glu	
530 535 540	
CCG CCC AAG AAA GGG CTG CTC CAG AGA CTC TTC AAG CGG CAG CAT CAG	1680
Pro Pro Lys Lys Gly Leu Leu Gln Arg Leu Phe Lys Arg Gln His Gln	
545 550 555 560	
AAC AAT TCC AAG AGT TCG CCC AGC TCC AAG ACC AGT TTT AAC CAC CAC	1728
Asn Asn Ser Lys Ser Ser Pro Ser Ser Lys Thr Ser Phe Asn His His	
565 570 575	
ATA AAC TCA AAC CAT GTC AGC TCG AAC TCC ACC GGA AGC AGC AGG GAT	1776
Ile Asn Ser Asn His Val Ser Ser Asn Ser Thr Gly Ser Ser Arg Asp	
580 585 590	
CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG	1824
Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly	
595 600 605	
GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG	1872
Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys	
610 615 620	
TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG	1920
Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu	
625 630 635 640	
ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC	1968
Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro	
645 650 655	
ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC	2016
Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr	
660 665 670	
CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA	2064
Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu	
675 680 685	
GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC	2112
Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr	
690 695 700	
AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC	2160
Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg	
705 710 715 720	
ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGC	2208

Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly	
725 730 735	
CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC	2256
His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala	
740 745 750	
GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC	2304
Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn	
755 760 765	
ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC	2352
Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr	
770 775 780	
CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC	2400
Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser	
785 790 795 800	
ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG	2448
Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met	
805 810 815	
GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC	2496
Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp	
820 825 830	
GAG CTG TAC AAG TAA	2511
Glu Leu Tyr Lys	
835	

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 836 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

Met Glu Leu Glu Asn Ile Val Ala Asn Thr Val Leu Leu Lys Ala Arg	
1 5 10 15	
Glu Gly Gly Gly Gly Lys Arg Lys Gly Lys Ser Lys Lys Trp Lys Glu	
20 25 30	
Ile Leu Lys Phe Pro His Ile Ser Gln Cys Glu Asp Leu Arg Arg Thr	
35 40 45	
Ile Asp Arg Asp Tyr Cys Ser Leu Cys Asp Lys Gln Pro Ile Gly Arg	
50 55 60	
Leu Leu Phe Arg Gln Phe Cys Glu Thr Arg Pro Gly Leu Glu Cys Tyr	
65 70 75 80	
Ile Gln Phe Leu Asp Ser Val Ala Glu Tyr Glu Val Thr Pro Asp Glu	
85 90 95	
Lys Leu Gly Glu Lys Gly Lys Glu Ile Met Thr Lys Tyr Leu Thr Pro	

			100					105				110				
Lys	Ser	Pro	Val	Phe	Ile	Ala	Gln	Val	Gly	Gln	Asp	Leu	Val	Ser	Gln	
		115					120					125				
Thr	Glu	Glu	Lys	Leu	Leu	Gln	Lys	Pro	Cys	Lys	Glu	Leu	Phe	Ser	Ala	
	130					135					140					
Cys	Ala	Gln	Ser	Val	His	Glu	Tyr	Leu	Arg	Gly	Glu	Pro	Phe	His	Glu	
145				150					155						160	
Tyr	Leu	Asp	Ser	Met	Phe	Phe	Asp	Arg	Phe	Leu	Gln	Trp	Lys	Trp	Leu	
			165						170					175		
Glu	Arg	Gln	Pro	Val	Thr	Lys	Asn	Thr	Phe	Arg	Gln	Tyr	Arg	Val	Leu	
			180					185					190			
Gly	Lys	Gly	Gly	Phe	Gly	Glu	Val	Cys	Ala	Cys	Gln	Val	Arg	Ala	Thr	
	195					200					205					
Gly	Lys	Met	Tyr	Ala	Cys	Lys	Arg	Leu	Glu	Lys	Lys	Arg	Ile	Lys	Lys	
	210					215					220					
Arg	Lys	Gly	Glu	Ser	Met	Ala	Leu	Asn	Glu	Lys	Gln	Ile	Leu	Glu	Lys	
225				230					235						240	
Val	Asn	Ser	Gln	Phe	Val	Val	Asn	Leu	Ala	Tyr	Ala	Tyr	Glu	Thr	Lys	
			245						250					255		
Asp	Ala	Leu	Cys	Leu	Val	Leu	Thr	Ile	Met	Asn	Gly	Gly	Asp	Leu	Lys	
	260							265					270			
Phe	His	Ile	Tyr	Asn	Met	Gly	Asn	Pro	Gly	Phe	Glu	Glu	Glu	Arg	Ala	
	275					280						285				
Leu	Phe	Tyr	Ala	Ala	Glu	Ile	Leu	Cys	Gly	Leu	Glu	Asp	Leu	His	Arg	
	290				295						300					
Glu	Asn	Thr	Val	Tyr	Arg	Asp	Leu	Lys	Pro	Glu	Asn	Ile	Leu	Leu	Asp	
305				310						315					320	
Asp	Tyr	Gly	His	Ile	Arg	Ile	Ser	Asp	Leu	Gly	Leu	Ala	Val	Lys	Ile	
			325					330						335		
Pro	Glu	Gly	Asp	Leu	Ile	Arg	Gly	Arg	Val	Gly	Thr	Val	Gly	Tyr	Met	
		340						345				350				
Ala	Pro	Glu	Val	Leu	Asn	Asn	Gln	Arg	Tyr	Gly	Leu	Ser	Pro	Asp	Tyr	
	355					360						365				
Trp	Gly	Leu	Gly	Cys	Leu	Ile	Tyr	Glu	Met	Ile	Glu	Gly	Gln	Ser	Pro	
	370				375						380					
Phe	Arg	Gly	Arg	Lys	Glu	Lys	Val	Lys	Arg	Glu	Glu	Val	Asp	Arg	Arg	
385				390						395					400	
Val	Leu	Glu	Thr	Glu	Glu	Val	Tyr	Ser	His	Lys	Phe	Ser	Glu	Glu	Ala	
			405					410						415		
Lys	Ser	Ile	Cys	Lys	Met	Leu	Leu	Thr	Lys	Asp	Ala	Lys	Gln	Arg	Leu	
		420						425					430			
Gly	Cys	Gln	Glu	Glu	Gly	Ala	Ala	Glu	Val	Lys	Arg	His	Pro	Phe	Phe	
	435					440			</							

	565		570		575
Ile Asn Ser Asn His Val Ser Ser Asn Ser Thr Gly Ser Ser Arg Asp					
580		585		590	
Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly					
595		600		605	
Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys					
610		615		620	
Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu					
625		630		635	640
Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro					
645		650		655	
Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr					
660		665		670	
Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu					
675		680		685	
Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr					
690		695		700	
Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg					
705		710		715	720
Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly					
725		730		735	
His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala					
740		745		750	
Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn					
755		760		765	
Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr					
770		775		780	
Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser					
785		790		795	800
Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met					
805		810		815	
Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp					
820		825		830	
Glu Leu Tyr Lys					
835					

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1893 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1890
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

ATG AGC AGA AGC AAG CGT GAC AAC AAT TTT TAT AGT GTA GAG ATT GGA	48
Met Ser Arg Ser Lys Arg Asp Asn Asn Phe Tyr Ser Val Glu Ile Gly	
1 5 10 15	
GAT TCT ACA TTC ACA GTC CTG AAA CGA TAT CAG AAT TTA AAA CCT ATA	96

Asp	Ser	Thr	Phe	Thr	Val	Leu	Lys	Arg	Tyr	Gln	Asn	Leu	Lys	Pro	Ile	
			20					25						30		
GGC	TCA	GGA	GCT	CAA	GGA	ATA	GTA	TGC	GCA	GCT	TAT	GAT	GCC	ATT	CTT	144
Gly	Ser	Gly	Ala	Gln	Gly	Ile	Val	Cys	Ala	Ala	Tyr	Asp	Ala	Ile	Leu	
		35					40					45				
GAA	AGA	AAT	GTT	GCA	ATC	AAG	AAG	CTA	AGC	CGA	CCA	TTT	CAG	AAT	CAG	192
Glu	Arg	Asn	Val	Ala	Ile	Lys	Lys	Leu	Ser	Arg	Pro	Phe	Gln	Asn	Gln	
		50				55					60					
ACT	CAT	GCC	AAG	CGG	GCC	TAC	AGA	GAG	CTA	GTT	CTT	ATG	AAA	TGT	GTT	240
Thr	His	Ala	Lys	Arg	Ala	Tyr	Arg	Glu	Leu	Val	Leu	Met	Lys	Cys	Val	
65					70				75					80		
AAT	CAC	AAA	AAT	ATA	ATT	GGC	CTT	TTG	AAT	GTT	TTC	ACA	CCA	CAG	AAA	288
Asn	His	Lys	Asn	Ile	Ile	Gly	Leu	Leu	Asn	Val	Phe	Thr	Pro	Gln	Lys	
				85					90					95		
TCC	CTA	GAA	GAA	TTT	CAA	GAT	GTT	TAC	ATA	GTC	ATG	GAG	CTC	ATG	GAT	336
Ser	Leu	Glu	Glu	Phe	Gln	Asp	Val	Tyr	Ile	Val	Met	Glu	Leu	Met	Asp	
			100					105					110			
GCA	AAT	CTT	TGC	CAA	GTG	ATT	CAG	ATG	GAG	CTA	GAT	CAT	GAA	AGA	ATG	384
Ala	Asn	Leu	Cys	Gln	Val	Ile	Gln	Met	Glu	Leu	Asp	His	Glu	Arg	Met	
		115					120					125				
TCC	TAC	CTT	CTC	TAT	CAG	ATG	CTG	TGT	GGA	ATC	AAG	CAC	CTT	CAT	TCT	432
Ser	Tyr	Leu	Leu	Tyr	Gln	Met	Leu	Cys	Gly	Ile	Lys	His	Leu	His	Ser	
		130				135					140					
GCT	GGA	ATT	ATT	CAT	CGG	GAC	TTA	AAG	CCC	AGT	AAT	ATA	GTA	GTA	AAA	480
Ala	Gly	Ile	Ile	His	Arg	Asp	Leu	Lys	Pro	Ser	Asn	Ile	Val	Val	Lys	
145				150						155					160	
TCT	GAT	TGC	ACT	TTG	AAG	ATT	CTT	GAC	TTC	GGT	CTG	GCC	AGG	ACT	GCA	528
Ser	Asp	Cys	Thr	Leu	Lys	Ile	Leu	Asp	Phe	Gly	Leu	Ala	Arg	Thr	Ala	
			165					170					175			
GGA	ACG	AGT	TTT	ATG	ATG	ACG	CCT	TAT	GTA	GTG	ACT	CGC	TAC	TAC	AGA	576
Gly	Thr	Ser	Phe	Met	Met	Thr	Pro	Tyr	Val	Val	Thr	Arg	Tyr	Tyr	Arg	
			180					185					190			
GCA	CCC	GAG	GTC	ATC	CTT	GGC	ATG	GGC	TAC	AAG	GAA	AAC	GTG	GAT	TTA	624
Ala	Pro	Glu	Val	Ile	Leu	Gly	Met	Gly	Tyr	Lys	Glu	Asn	Val	Asp	Leu	
		195					200					205				
TGG	TCT	GTG	GGG	TGC	ATT	ATG	GGA	GAA	ATG	GTT	TGC	CAC	AAA	ATC	CTC	672
Trp	Ser	Val	Gly	Cys	Ile	Met	Gly	Glu	Met	Val	Cys	His	Lys	Ile	Leu	
		210				215					220					
TTT	CCA	GGA	AGG	GAC	TAT	ATT	GAT	CAG	TGG	AAT	AAA	GTT	ATT	GAA	CAG	720
Phe	Pro	Gly	Arg	Asp	Tyr	Ile	Asp	Gln	Trp	Asn	Lys	Val	Ile	Glu	Gln	
225				230						235				240		
CTT	GGA	ACA	CCA	TGT	CCT	GAA	TTC	ATG	AAG	AAA	CTG	CAA	CCA	ACA	GTA	768
Leu	Gly	Thr	Pro	Cys	Pro	Glu	Phe	Met	Lys	Lys	Leu	Gln	Pro	Thr	Val	
				245					250					255		

AGG ACT TAC GTT GAA AAC AGA CCT AAA TAT GCT GGA TAT AGC TTT GAG	816
Arg Thr Tyr Val Glu Asn Arg Pro Lys Tyr Ala Gly Tyr Ser Phe Glu	
260 265 270	
AAA CTC TTC CCT GAT GTC CTT TTC CCA GCT GAC TCA GAA CAC AAC AAA	864
Lys Leu Phe Pro Asp Val Leu Phe Pro Ala Asp Ser Glu His Asn Lys	
275 280 285	
CTT AAA GCC AGT CAG GCA AGG GAT TTG TTA TCC AAA ATG CTG GTA ATA	912
Leu Lys Ala Ser Gln Ala Arg Asp Leu Leu Ser Lys Met Leu Val Ile	
290 295 300	
GAT GCA TCT AAA AGG ATC TCT GTA GAT GAA GCT CTC CAA CAC CCG TAC	960
Asp Ala Ser Lys Arg Ile Ser Val Asp Glu Ala Leu Gln His Pro Tyr	
305 310 315 320	
ATC AAT GTC TGG TAT GAT CCT TCT GAA GCA GAA GCT CCA CCA CCA AAG	1008
Ile Asn Val Trp Tyr Asp Pro Ser Glu Ala Glu Ala Pro Pro Pro Lys	
325 330 335	
ATC CCT GAC AAG CAG TTA GAT GAA AGG GAA CAC ACA ATA GAA GAG TGG	1056
Ile Pro Asp Lys Gln Leu Asp Glu Arg Glu His Thr Ile Glu Glu Trp	
340 345 350	
AAA GAA TTG ATA TAT AAG GAA GTT ATG GAC TTG GAG GAG AGA ACC AAG	1104
Lys Glu Leu Ile Tyr Lys Glu Val Met Asp Leu Glu Glu Arg Thr Lys	
355 360 365	
AAT GGA GTT ATA CGG GGG CAG CCC TCT CCT TTA GCA CAG GTG CAG CAG	1152
Asn Gly Val Ile Arg Gly Gln Pro Ser Pro Leu Ala Gln Val Gln Gln	
370 375 380	
TGG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC	1200
Trp Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe	
385 390 395 400	
ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC	1248
Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly	
405 410 415	
CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC	1296
His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly	
420 425 430	
AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC	1344
Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro	
435 440 445	
TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC	1392
Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser	
450 455 460	
CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG	1440
Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met	
465 470 475 480	
CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC	1488

Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly	
485 490 495	
AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG	1536
Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val	
500 505 510	
AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC	1584
Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile	
515 520 525	
CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC	1632
Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile	
530 535 540	
ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC	1680
Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg	
545 550 555 560	
CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG	1728
His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln	
565 570 575	
AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC	1776
Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr	
580 585 590	
CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT	1824
Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp	
595 600 605	
CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC	1872
His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly	
610 615 620	
ATG GAC GAG CTG TAC AAG TAA	1893
Met Asp Glu Leu Tyr Lys	
625 630	

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 630 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Met Ser Arg Ser Lys Arg Asp Asn Asn Phe Tyr Ser Val Glu Ile Gly	
1 5 10 15	
Asp Ser Thr Phe Thr Val Leu Lys Arg Tyr Gln Asn Leu Lys Pro Ile	
20 25 30	
Gly Ser Gly Ala Gln Gly Ile Val Cys Ala Ala Tyr Asp Ala Ile Leu	

35					40					45				
Glu	Arg	Asn	Val	Ala	Ile	Lys	Leu	Ser	Arg	Pro	Phe	Gln	Asn	Gln
50					55					60				
Thr	His	Ala	Lys	Arg	Ala	Tyr	Arg	Glu	Leu	Val	Leu	Met	Lys	Cys
65					70					75				80
Asn	His	Lys	Asn	Ile	Ile	Gly	Leu	Leu	Asn	Val	Phe	Thr	Pro	Gln
				85					90					95
Ser	Leu	Glu	Glu	Phe	Gln	Asp	Val	Tyr	Ile	Val	Met	Glu	Leu	Met
				100				105					110	Asp
Ala	Asn	Leu	Cys	Gln	Val	Ile	Gln	Met	Glu	Leu	Asp	His	Glu	Arg
				115				120					125	Met
Ser	Tyr	Leu	Leu	Tyr	Gln	Met	Leu	Cys	Gly	Ile	Lys	His	Leu	His
				130				135					140	Ser
Ala	Gly	Ile	Ile	His	Arg	Asp	Leu	Lys	Pro	Ser	Asn	Ile	Val	Val
145					150					155				160
Ser	Asp	Cys	Thr	Leu	Lys	Ile	Leu	Asp	Phe	Gly	Leu	Ala	Arg	Thr
				165					170					175
Gly	Thr	Ser	Phe	Met	Met	Thr	Pro	Tyr	Val	Val	Thr	Arg	Tyr	Arg
				180				185					190	
Ala	Pro	Glu	Val	Ile	Leu	Gly	Met	Gly	Tyr	Lys	Glu	Asn	Val	Asp
				195				200					205	Leu
Trp	Ser	Val	Gly	Cys	Ile	Met	Gly	Glu	Met	Val	Cys	His	Lys	Ile
				210				215					220	Leu
Phe	Pro	Gly	Arg	Asp	Tyr	Ile	Asp	Gln	Trp	Asn	Lys	Val	Ile	Glu
225					230					235				240
Leu	Gly	Thr	Pro	Cys	Pro	Glu	Phe	Met	Lys	Lys	Leu	Gln	Pro	Thr
				245					250					255
Arg	Thr	Tyr	Val	Glu	Asn	Arg	Pro	Lys	Tyr	Ala	Gly	Tyr	Ser	Phe
				260				265						270
Lys	Leu	Phe	Pro	Asp	Val	Leu	Phe	Pro	Ala	Asp	Ser	Glu	His	Asn
				275				280					285	Lys
Leu	Lys	Ala	Ser	Gln	Ala	Arg	Asp	Leu	Leu	Ser	Lys	Met	Leu	Val
				290				295					300	Ile
Asp	Ala	Ser	Lys	Arg	Ile	Ser	Val	Asp	Glu	Ala	Leu	Gln	His	Pro
305					310					315				320
Ile	Asn	Val	Trp	Tyr	Asp	Pro	Ser	Glu	Ala	Glu	Ala	Pro	Pro	Pro
				325					330					335
Ile	Pro	Asp	Lys	Gln	Leu	Asp	Glu	Arg	Glu	His	Thr	Ile	Glu	Glu
				340				345					350	Trp
Lys	Glu	Leu	Ile	Tyr	Lys	Glu	Val	Met	Asp	Leu	Glu	Glu	Arg	Thr
				355				360					365	Lys
Asn	Gly	Val	Ile	Arg	Gly	Gln	Pro	Ser	Pro	Leu	Ala	Gln	Val	Gln
				370				375					380	Gln
Trp	Asp	Pro	Pro	Val	Ala	Thr	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu
385					390					395				400
Thr	Gly	Val	Val	Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn
				405					410					415
His	Lys	Phe	Ser	Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr
				420				425					430	Gly
Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val
				435				440					445	Pro
Trp	Pro	Thr	Leu	Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe
				450				455					460	Ser
Arg	Tyr	Pro	Asp	His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala
465					470					475				480
Pro	Glu	Gly	Tyr	Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp
				485					490					495
Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu
														Val

```

          500          505          510
Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile
          515          520          525
Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile
          530          535          540
Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg
          545          550          555          560
His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln
          565          570          575
Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr
          580          585          590
Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp
          595          600          605
His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly
          610          615          620
Met Asp Glu Leu Tyr Lys
          625          630

```

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1821 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1818
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

```

ATG TCT CAG GAG AGG CCC ACG TTC TAC CGG CAG GAG CTG AAC AAG ACA      48
Met Ser Gln Glu Arg Pro Thr Phe Tyr Arg Gln Glu Leu Asn Lys Thr
  1              5              10              15

ATC TGG GAG GTG CCC GAG CGT TAC CAG AAC CTG TCT CCA GTG GGC TCT      96
Ile Trp Glu Val Pro Glu Arg Tyr Gln Asn Leu Ser Pro Val Gly Ser
  20              25              30

GGC GCC TAT GGC TCT GTG TGT GCT GCT TTT GAC ACA AAA ACG GGG TTA      144
Gly Ala Tyr Gly Ser Val Cys Ala Ala Phe Asp Thr Lys Thr Gly Leu
  35              40              45

CGT GTG GCA GTG AAG AAG CTC TCC AGA CCA TTT CAG TCC ATC ATT CAT      192
Arg Val Ala Val Lys Lys Leu Ser Arg Pro Phe Gln Ser Ile Ile His
  50              55              60

GGG AAA AGA ACC TAC AGA GAA CTG CGG TTA CTT AAA CAT ATG AAA CAT      240
Ala Lys Arg Thr Tyr Arg Glu Leu Arg Leu Leu Lys His Met Lys His
  65              70              75              80

GAA AAT GTG ATT GGT CTG TTG GAC GTT TTT ACA CCT GCA AGG TCT CTG      288
Glu Asn Val Ile Gly Leu Leu Asp Val Phe Thr Pro Ala Arg Ser Leu
  85              90              95

```

GAG GAA TTC AAT GAT GTG TAT CTG GTG ACC CAT CTC ATG GGG GCA GAT Glu Glu Phe Asn Asp Val Tyr Leu Val Thr His Leu Met Gly Ala Asp 100 105 110	336
CTG AAC AAC ATT GTG AAA TGT CAG AAG CTT ACA GAT GAC CAT GTT CAG Leu Asn Asn Ile Val Lys Cys Gln Lys Leu Thr Asp Asp His Val Gln 115 120 125	384
TTC CTT ATC TAC CAA ATT CTC CGA GGT CTA AAG TAT ATA CAT TCA GCT Phe Leu Ile Tyr Gln Ile Leu Arg Gly Leu Lys Tyr Ile His Ser Ala 130 135 140	432
GAC ATA ATT CAC AGG GAC CTA AAA CCT AGT AAT CTA GCT GTG AAT GAA Asp Ile Ile His Arg Asp Leu Lys Pro Ser Asn Leu Ala Val Asn Glu 145 150 155 160	480
GAC TGT GAG CTG AAG ATT CTG GAT TTT GGA CTG GCT CGG CAC ACA GAT Asp Cys Glu Leu Lys Ile Leu Asp Phe Gly Leu Ala Arg His Thr Asp 165 170 175	528
GAT GAA ATG ACA GGC TAC GTG GCC ACT AGG TGG TAC AGG GCT CCT GAG Asp Glu Met Thr Gly Tyr Val Ala Thr Arg Trp Tyr Arg Ala Pro Glu 180 185 190	576
ATC ATG CTG AAC TGG ATG CAT TAC AAC CAG ACA GTT GAT ATT TGG TCA Ile Met Leu Asn Trp Met His Tyr Asn Gln Thr Val Asp Ile Trp Ser 195 200 205	624
GTG GGA TGC ATA ATG GCC GAG CTG TTG ACT GGA AGA ACA TTG TTT CCT Val Gly Cys Ile Met Ala Glu Leu Leu Thr Gly Arg Thr Leu Phe Pro 210 215 220	672
GGT ACA GAC CAT ATT GAT CAG TTG AAG CTC ATT TTA AGA CTC GTT GGA Gly Thr Asp His Ile Asp Gln Leu Lys Leu Ile Leu Arg Leu Val Gly 225 230 235 240	720
ACC CCA GGG GCT GAG CTT TTG AAG AAA ATC TCC TCA GAG TCT GCA AGA Thr Pro Gly Ala Glu Leu Leu Lys Lys Ile Ser Ser Glu Ser Ala Arg 245 250 255	768
AAC TAT ATT CAG TCT TTG ACT CAG ATG CCG AAG ATG AAC TTT GCG AAT Asn Tyr Ile Gln Ser Leu Thr Gln Met Pro Lys Met Asn Phe Ala Asn 260 265 270	816
GTA TTT ATT GGT GCC AAT CCC CTG GCT GTC GAC TTG CTG GAG AAG ATG Val Phe Ile Gly Ala Asn Pro Leu Ala Val Asp Leu Leu Glu Lys Met 275 280 285	864
CTT GTA TTG GAC TCA GAT AAG AGA ATT ACA GCG GCC CAA GCC CTT GCA Leu Val Leu Asp Ser Asp Lys Arg Ile Thr Ala Ala Gln Ala Leu Ala 290 295 300	912
CAT GCC TAC TTT GCT CAG TAC CAC GAT CCT GAT GAT GAA CCA GTG GCC His Ala Tyr Phe Ala Gln Tyr His Asp Pro Asp Asp Glu Pro Val Ala 305 310 315 320	960
GAT CCT TAT GAT CAG TCC TTT GAA AGC AGG GAC CTC CTT ATA GAT GAG	1008

Asp Pro Tyr Asp Gln Ser Phe Glu Ser Arg Asp Leu Leu Ile Asp Glu	
325 330 335	
TGG AAA AGC CTG ACC TAT GAT GAA GTC ATC AGC TTT GTG CCA CCA CCC	1056
Trp Lys Ser Leu Thr Tyr Asp Glu Val Ile Ser Phe Val Pro Pro Pro	
340 345 350	
CTT GAC CAA GAA GAG ATG GAG TCC GAG GAT CCA CCG GTC GCC ACC ATG	1104
Leu Asp Gln Glu Glu Met Glu Ser Glu Asp Pro Pro Val Ala Thr Met	
355 360 365	
GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC	1152
Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val	
370 375 380	
GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG	1200
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu	
385 390 395 400	
GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC	1248
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys	
405 410 415	
ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG	1296
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu	
420 425 430	
ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG	1344
Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln	
435 440 445	
CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC	1392
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg	
450 455 460	
ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG	1440
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val	
465 470 475 480	
AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC	1488
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile	
485 490 495	
GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC	1536
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn	
500 505 510	
TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC	1584
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly	
515 520 525	
ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG	1632
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val	
530 535 540	
CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC	1680
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro	
545 550 555 560	

GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC 1728
 Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
 565 570 575

AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG 1776
 Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
 580 585 590

ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA 1821
 Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
 595 600 605

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 606 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Met	Ser	Gln	Glu	Arg	Pro	Thr	Phe	Tyr	Arg	Gln	Glu	Leu	Asn	Lys	Thr
1				5					10					15	
Ile	Trp	Glu	Val	Pro	Glu	Arg	Tyr	Gln	Asn	Leu	Ser	Pro	Val	Gly	Ser
		20						25					30		
Gly	Ala	Tyr	Gly	Ser	Val	Cys	Ala	Phe	Asp	Thr	Lys	Thr	Gly	Leu	
		35					40				45				
Arg	Val	Ala	Val	Lys	Lys	Leu	Ser	Arg	Pro	Phe	Gln	Ser	Ile	Ile	His
		50				55					60				
Ala	Lys	Arg	Thr	Tyr	Arg	Glu	Leu	Arg	Leu	Leu	Lys	His	Met	Lys	His
65					70					75				80	
Glu	Asn	Val	Ile	Gly	Leu	Leu	Asp	Val	Phe	Thr	Pro	Ala	Arg	Ser	Leu
				85				90						95	
Glu	Glu	Phe	Asn	Asp	Val	Tyr	Leu	Val	Thr	His	Leu	Met	Gly	Ala	Asp
				100				105					110		
Leu	Asn	Asn	Ile	Val	Lys	Cys	Gln	Lys	Leu	Thr	Asp	Asp	His	Val	Gln
				115				120					125		
Phe	Leu	Ile	Tyr	Gln	Ile	Leu	Arg	Gly	Leu	Lys	Tyr	Ile	His	Ser	Ala
				130				135					140		
Asp	Ile	Ile	His	Arg	Asp	Leu	Lys	Pro	Ser	Asn	Leu	Ala	Val	Asn	Glu
145					150					155				160	
Asp	Cys	Glu	Leu	Lys	Ile	Leu	Asp	Phe	Gly	Leu	Ala	Arg	His	Thr	Asp
				165					170					175	
Asp	Glu	Met	Thr	Gly	Tyr	Val	Ala	Thr	Arg	Trp	Tyr	Arg	Ala	Pro	Glu
				180				185					190		
Ile	Met	Leu	Asn	Trp	Met	His	Tyr	Asn	Gln	Thr	Val	Asp	Ile	Trp	Ser
				195				200					205		
Val	Gly	Cys	Ile	Met	Ala	Glu	Leu	Leu	Thr	Gly	Arg	Thr	Leu	Phe	Pro
				210				215					220		
Gly	Thr	Asp	His	Ile	Asp	Gln	Leu	Lys	Leu	Ile	Leu	Arg	Leu	Val	Gly
225					230					235				240	
Thr	Pro	Gly	Ala	Glu	Leu	Leu	Lys	Lys	Ile	Ser	Ser	Glu	Ser	Ala	Arg

					245					250					255	
Asn	Tyr	Ile	Gln	Ser	Leu	Thr	Gln	Met	Pro	Lys	Met	Asn	Phe	Ala	Asn	
			260					265						270		
Val	Phe	Ile	Gly	Ala	Asn	Pro	Leu	Ala	Val	Asp	Leu	Leu	Glu	Lys	Met	
		275					280					285				
Leu	Val	Leu	Asp	Ser	Asp	Lys	Arg	Ile	Thr	Ala	Ala	Gln	Ala	Leu	Ala	
	290					295					300					
His	Ala	Tyr	Phe	Ala	Gln	Tyr	His	Asp	Pro	Asp	Asp	Glu	Pro	Val	Ala	
305					310					315					320	
Asp	Pro	Tyr	Asp	Gln	Ser	Phe	Glu	Ser	Arg	Asp	Leu	Leu	Ile	Asp	Glu	
				325					330					335		
Trp	Lys	Ser	Leu	Thr	Tyr	Asp	Glu	Val	Ile	Ser	Phe	Val	Pro	Pro	Pro	
			340					345					350			
Leu	Asp	Gln	Glu	Glu	Met	Glu	Ser	Glu	Asp	Pro	Pro	Val	Ala	Thr	Met	
		355					360					365				
Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	Val	
	370					375					380					
Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	Glu	
385					390					395					400	
Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	
				405					410					415		
Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	Leu	
		420						425					430			
Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	Gln	
	435					440						445				
His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	Arg	
450						455					460					
Thr	Ile	Phe	Phe	Lys	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val		
465				470					475					480		
Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile	
				485					490					495		
Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	Asn	
		500						505					510			
Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	Gly	
	515						520					525				
Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	Val	
	530					535					540					
Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	Pro	
545				550						555					560	
Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	Ser	
				565					570					575		
Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	Val	
		580														

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2913 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence

(B) LOCATION: 1...2910

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATG ACT GCT GAG GGG TAC CAG TAC AGA GCG CTG TAT GAT TAT AAA AAG	48
Met Ser Ala Glu Gly Tyr Gln Tyr Arg Ala Leu Tyr Asp Tyr Lys Lys	
1 5 10 15	
GAA AGA GAA GAA GAT ATT GAC TTG CAC TTG GGT GAC ATA TTG ACT GTG	96
Glu Arg Glu Glu Asp Ile Asp Leu His Leu Gly Asp Ile Leu Thr Val	
20 25 30	
AAT AAA GGG TCC TTA GTA GCT CTT GGA TTC AGT GAT GGA CAG GAA GCC	144
Asn Lys Gly Ser Leu Val Ala Leu Gly Phe Ser Asp Gly Gln Glu Ala	
35 40 45	
AGG CCT GAA GAA ATT GGC TGG TTA AAT GGC TAT AAT GAA ACC ACA GGG	192
Arg Pro Glu Glu Ile Gly Trp Leu Asn Gly Tyr Asn Glu Thr Thr Gly	
50 55 60	
GAA AGG GGG GAC TTT CCG GGA ACT TAC GTA GAA TAT ATT GGA AGG AAA	240
Glu Arg Gly Asp Phe Pro Gly Thr Tyr Val Glu Tyr Ile Gly Arg Lys	
65 70 75 80	
AAA ATC TCG CCT CCC ACA CCA AAG CCC CGG CCA CCT CGG CCT CTT CCT	288
Lys Ile Ser Pro Pro Thr Pro Lys Pro Arg Pro Pro Arg Pro Leu Pro	
85 90 95	
GTT GCA CCA GGT TCT TCG AAA ACT GAA GCA GAT GTT GAA CAA CAA GCT	336
Val Ala Pro Gly Ser Ser Lys Thr Glu Ala Asp Val Glu Gln Gln Ala	
100 105 110	
TTG ACT CTC CCG GAT CTT GCA GAG CAG TTT GCC CCT CCT GAC ATT GCC	384
Leu Thr Leu Pro Asp Leu Ala Glu Gln Phe Ala Pro Pro Asp Ile Ala	
115 120 125	
CCG CCT CTT CTT ATC AAG CTC GTG GAA GCC ATT GAA AAG AAA GGT CTG	432
Pro Pro Leu Leu Ile Lys Leu Val Glu Ala Ile Glu Lys Lys Gly Leu	
130 135 140	
GAA TGT TCA ACT CTA TAC AGA ACA CAG AGC TCC AGC AAC CTG GCA GAA	480
Glu Cys Ser Thr Leu Tyr Arg Thr Gln Ser Ser Ser Asn Leu Ala Glu	
145 150 155 160	
TTA CGA CAG CTT CTT GAT TGT GAT ACA CCC TCC GTG GAC TTG GAA ATG	528
Leu Arg Gln Leu Leu Asp Cys Asp Thr Pro Ser Val Asp Leu Glu Met	
165 170 175	
ATC GAT GTG CAC GTT TTG GCT GAC GCT TTC AAA CGC TAT CTC CTG GAC	576
Ile Asp Val His Val Leu Ala Asp Ala Phe Lys Arg Tyr Leu Leu Asp	
180 185 190	
TTA CCA AAT CCT GTC ATT CCA GCA GCC GTT TAC AGT GAA ATG ATT TCT	624
Leu Pro Asn Pro Val Ile Pro Ala Ala Val Tyr Ser Glu Met Ile Ser	
195 200 205	
TTA GCT CCA GAA GTA CAA AGC TCC GAA GAA TAT ATT CAG CTA TTG AAG	672

Leu Ala Pro Glu Val Gln Ser Ser Glu Glu Tyr Ile Gln Leu Leu Lys	
210	220
AAG CTT ATT AGG TCG CCT AGC ATA CCT CAT CAG TAT TGG CTT ACG CTT	720
Lys Leu Ile Arg Ser Pro Ser Ile Pro His Gln Tyr Trp Leu Thr Leu	
225	240
CAG TAT TTG TTA AAA CAT TTC TTC AAG CTC TCT CAA ACC TCC AGC AAA	768
Gln Tyr Leu Leu Lys His Phe Phe Lys Leu Ser Gln Thr Ser Ser Lys	
245	255
AAT CTG TTG AAT GCA AGA GTA CTC TCT GAA ATT TTC AGC CCT ATG CTT	816
Asn Leu Leu Asn Ala Arg Val Leu Ser Glu Ile Phe Ser Pro Met Leu	
260	270
TTC AGA TTC TCA GCA GCC AGC TCT GAT AAT ACT GAA AAC CTC ATA AAA	864
Phe Arg Phe Ser Ala Ala Ser Ser Asp Asn Thr Glu Asn Leu Ile Lys	
275	285
GTT ATA GAA ATT TTA ATC TCA ACT GAA TGG AAT GAA CGA CAG CCT GCA	912
Val Ile Glu Ile Leu Ile Ser Thr Glu Trp Asn Glu Arg Gln Pro Ala	
290	300
CCA GCA CTG CCT CCT AAA CCA CCA AAA CCT ACT ACT GTA GCC AAC AAC	960
Pro Ala Leu Pro Pro Lys Pro Pro Lys Pro Thr Thr Val Ala Asn Asn	
305	320
GGT ATG AAT AAC AAT ATG TCC TTA CAA AAT GCT GAA TGG TAC TGG GGA	1008
Gly Met Asn Asn Asn Met Ser Leu Gln Asn Ala Glu Trp Tyr Trp Gly	
325	335
GAT ATC TCG AGG GAA GAA GTG AAT GAA AAA CTT CGA GAT ACA GCA GAC	1056
Asp Ile Ser Arg Glu Glu Val Asn Glu Lys Leu Arg Asp Thr Ala Asp	
340	350
GGG ACC TTT TTG GTA CGA GAT GCG TCT ACT AAA ATG CAT GGT GAT TAT	1104
Gly Thr Phe Leu Val Arg Asp Ala Ser Thr Lys Met His Gly Asp Tyr	
355	365
ACT CTT ACA CTA AGG AAA GGG GGA AAT AAC AAA TTA ATC AAA ATA TTT	1152
Thr Leu Thr Leu Arg Lys Gly Gly Asn Asn Lys Leu Ile Lys Ile Phe	
370	380
CAT CGA GAT GGG AAA TAT GGC TTC TCT GAC CCA TTA ACC TTC AGT TCT	1200
His Arg Asp Gly Lys Tyr Gly Phe Ser Asp Pro Leu Thr Phe Ser Ser	
385	400
GTG GTT GAA TTA ATA AAC CAC TAC CGG AAT GAA TCT CTA GCT CAG TAT	1248
Val Val Glu Leu Ile Asn His Tyr Arg Asn Glu Ser Leu Ala Gln Tyr	
405	415
AAT CCC AAA TTG GAT GTG AAA TTA CTT TAT CCA GTA TCC AAA TAC CAA	1296
Asn Pro Lys Leu Asp Val Lys Leu Leu Tyr Pro Val Ser Lys Tyr Gln	
420	430
CAG GAT CAA GTT GTC AAA GAA GAT AAT ATT GAA GCT GTA GGG AAA AAA	1344
Gln Asp Gln Val Val Lys Glu Asp Asn Ile Glu Ala Val Gly Lys Lys	
435	445

TTA CAT GAA TAT AAC ACT CAG TTT CAA GAA AAA AGT CGA GAA TAT GAT Leu His Glu Tyr Asn Thr Gln Phe Gln Glu Lys Ser Arg Glu Tyr Asp 450 455 460	1392
AGA TTA TAT GAA GAA TAT ACC CGC ACA TCC CAG GAA ATC CAA ATG AAA Arg Leu Tyr Glu Glu Tyr Thr Arg Thr Ser Gln Glu Ile Gln Met Lys 465 470 475 480	1440
AGG ACA GCT ATT GAA GCA TTT AAT GAA ACC ATA AAA ATA TTT GAA GAA Arg Thr Ala Ile Glu Ala Phe Asn Glu Thr Ile Lys Ile Phe Glu Glu 485 490 495	1488
CAG TGC CAG ACC CAA GAG CGG TAC AGC AAA GAA TAC ATA GAA AAG TTT Gln Cys Gln Thr Gln Glu Arg Tyr Ser Lys Glu Tyr Ile Glu Lys Phe 500 505 510	1536
AAA CGT GAA GGC AAT GAG AAA GAA ATA CAA AGG ATT ATG CAT AAT TAT Lys Arg Glu Gly Asn Glu Lys Glu Ile Gln Arg Ile Met His Asn Tyr 515 520 525	1584
GAT AAG TTG AAG TCT CGA ATC AGT GAA ATT ATT GAC ACT AGA AGA AGA Asp Lys Leu Lys Ser Arg Ile Ser Glu Ile Ile Asp Ser Arg Arg Arg 530 535 540	1632
TTG GAA GAA GAC TTG AAG AAG CAG GCA GCT GAG TAT CGA GAA ATT GAC Leu Glu Glu Asp Leu Lys Lys Gln Ala Ala Glu Tyr Arg Glu Ile Asp 545 550 555 560	1680
AAA CGT ATG AAC AGC ATT AAA CCA GAC CTT ATC CAG CTG AGA AAG ACG Lys Arg Met Asn Ser Ile Lys Pro Asp Leu Ile Gln Leu Arg Lys Thr 565 570 575	1728
AGA GAC CAA TAC TTG ATG TGG TTG ACT CAA AAA GGT GTT CGG CAA AAG Arg Asp Gln Tyr Leu Met Trp Leu Thr Gln Lys Gly Val Arg Gln Lys 580 585 590	1776
AAG TTG AAC GAG TGG TTG GGC AAT GAA AAC ACT GAA GAC CAA TAT TCA Lys Leu Asn Glu Trp Leu Gly Asn Glu Asn Thr Glu Asp Gln Tyr Ser 595 600 605	1824
CTG GTG GAA GAT GAT GAA GAT TTG CCC CAT CAT GAT GAG AAG ACA TGG Leu Val Glu Asp Asp Glu Asp Leu Pro His His Asp Glu Lys Thr Trp 610 615 620	1872
AAT GTT GGA AGC AGC AAC CGA AAC AAA GCT GAA AAC CTG TTG CGA GGG Asn Val Gly Ser Ser Asn Arg Asn Lys Ala Glu Asn Leu Leu Arg Gly 625 630 635 640	1920
AAG CGA GAT GGC ACT TTT CTT GTC CGG GAG AGC AGT AAA CAG GGC TGC Lys Arg Asp Gly Thr Phe Leu Val Arg Glu Ser Ser Lys Gln Gly Cys 645 650 655	1968
TAT GCC TGC TCT GTA GTG GTG GAC GGC GAA GTA AAG CAT TGT GTC ATA Tyr Ala Cys Ser Val Val Val Asp Gly Glu Val Lys His Cys Val Ile 660 665 670	2016
AAC AAA ACA GCA ACT GGC TAT GGC TTT GCC GAG CCC TAT AAC TTG TAC	2064

Asn	Lys	Thr	Ala	Thr	Gly	Tyr	Gly	Phe	Ala	Glu	Pro	Tyr	Asn	Leu	Tyr	
		675					680					685				
AGC	TCT	CTG	AAA	GAA	CTG	GTG	CTA	CAT	TAC	CAA	CAC	ACC	TCC	CTT	GTG	2112
Ser	Ser	Leu	Lys	Glu	Leu	Val	Leu	His	Tyr	Gln	His	Thr	Ser	Leu	Val	
		690				695				700						
CAG	CAC	AAC	GAC	TCC	CTC	AAT	GTC	ACA	CTA	GCC	TAC	CCA	GTA	TAT	GCA	2160
Gln	His	Asn	Asp	Ser	Leu	Asn	Val	Thr	Leu	Ala	Tyr	Pro	Val	Tyr	Ala	
705					710				715					720		
CAG	CAG	AGG	CGA	CAG	GAT	CCA	CCG	GTC	GCC	ACC	ATG	GTG	AGC	AAG	GGC	2208
Gln	Gln	Arg	Arg	Gln	Asp	Pro	Pro	Val	Ala	Thr	Met	Val	Ser	Lys	Gly	
				725					730					735		
GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	GTC	GAG	CTG	GAC	GGC	2256
Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	
			740					745					750			
GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	GAG	GGC	GAG	GGC	GAT	2304
Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	
		755				760					765					
GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	AAG	TTC	ATC	TGC	ACC	ACC	GGC	AAG	2352
Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	
		770				775					780					
CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	CTG	ACC	TAC	GGC	GTG	2400
Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	
785					790				795					800		
CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	CAC	ATG	AAG	CAG	CAC	GAC	TTC	TTC	2448
Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	Gln	His	Asp	Phe	Phe	
				805				810						815		
AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	CGC	ACC	ATC	TTC	TTC	2496
Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	
			820					825					830			
AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	GTG	AAG	TTC	GAG	GGC	2544
Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	
			835				840					845				
GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	CTG	AAG	GGC	ATC	GAC	TTC	AAG	GAG	2592
Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	
		850				855					860					
GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	CTG	GAG	TAC	AAC	TAC	AAC	AGC	CAC	2640
Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	
				865		870				875				880		
AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	GGC	ATC	AAG	GTG	AAC	2688
Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	
				885				890						895		
TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	GTG	CAG	CTC	GCC	GAC	2736
Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	
			900					905						910		

CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC 2784
 His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro
 915 920 925
 GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC 2832
 Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn
 930 935 940
 GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG 2880
 Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly
 945 950 955 960
 ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA 2913
 Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
 965 970

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 970 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Met Ser Ala Glu Gly Tyr Gln Tyr Arg Ala Leu Tyr Asp Tyr Lys Lys
 1 5 10 15
 Glu Arg Glu Glu Asp Ile Asp Leu His Leu Gly Asp Ile Leu Thr Val
 20 25 30
 Asn Lys Gly Ser Leu Val Ala Leu Gly Phe Ser Asp Gly Gln Glu Ala
 35 40 45
 Arg Pro Glu Glu Ile Gly Trp Leu Asn Gly Tyr Asn Glu Thr Thr Gly
 50 55 60
 Glu Arg Gly Asp Phe Pro Gly Thr Tyr Val Glu Tyr Ile Gly Arg Lys
 65 70 75 80
 Lys Ile Ser Pro Pro Thr Pro Lys Pro Arg Pro Pro Arg Pro Leu Pro
 85 90 95
 Val Ala Pro Gly Ser Ser Lys Thr Glu Ala Asp Val Glu Gln Gln Ala
 100 105 110
 Leu Thr Leu Pro Asp Leu Ala Glu Gln Phe Ala Pro Pro Asp Ile Ala
 115 120 125
 Pro Pro Leu Leu Ile Lys Leu Val Glu Ala Ile Glu Lys Lys Gly Leu
 130 135 140
 Glu Cys Ser Thr Leu Tyr Arg Thr Gln Ser Ser Ser Asn Leu Ala Glu
 145 150 155 160
 Leu Arg Gln Leu Leu Asp Cys Asp Thr Pro Ser Val Asp Leu Glu Met
 165 170 175
 Ile Asp Val His Val Leu Ala Asp Ala Phe Lys Arg Tyr Leu Leu Asp
 180 185 190
 Leu Pro Asn Pro Val Ile Pro Ala Ala Val Tyr Ser Glu Met Ile Ser
 195 200 205
 Leu Ala Pro Glu Val Gln Ser Ser Glu Glu Tyr Ile Gln Leu Leu Lys

210		215		220
Lys Leu Ile Arg Ser Pro Ser Ile Pro His Gln Tyr Trp Leu Thr Leu				
225		230		235
Gln Tyr Leu Leu Lys His Phe Phe Lys Leu Ser Gln Thr Ser Ser Lys				240
	245		250	255
Asn Leu Leu Asn Ala Arg Val Leu Ser Glu Ile Phe Ser Pro Met Leu				
	260		265	270
Phe Arg Phe Ser Ala Ala Ser Ser Asp Asn Thr Glu Asn Leu Ile Lys				
	275		280	285
Val Ile Glu Ile Leu Ile Ser Thr Glu Trp Asn Glu Arg Gln Pro Ala				
	290		295	300
Pro Ala Leu Pro Pro Lys Pro Pro Lys Pro Thr Thr Val Ala Asn Asn				
305		310		315
Gly Met Asn Asn Asn Met Ser Leu Gln Asn Ala Glu Trp Tyr Trp Gly				320
	325		330	335
Asp Ile Ser Arg Glu Glu Val Asn Glu Lys Leu Arg Asp Thr Ala Asp				
	340		345	350
Gly Thr Phe Leu Val Arg Asp Ala Ser Thr Lys Met His Gly Asp Tyr				
	355		360	365
Thr Leu Thr Leu Arg Lys Gly Gly Asn Asn Lys Leu Ile Lys Ile Phe				
	370		375	380
His Arg Asp Gly Lys Tyr Gly Phe Ser Asp Pro Leu Thr Phe Ser Ser				
385		390		395
Val Val Glu Leu Ile Asn His Tyr Arg Asn Glu Ser Leu Ala Gln Tyr				
	405		410	415
Asn Pro Lys Leu Asp Val Lys Leu Leu Tyr Pro Val Ser Lys Tyr Gln				
	420		425	430
Gln Asp Gln Val Val Lys Glu Asp Asn Ile Glu Ala Val Gly Lys Lys				
	435		440	445
Leu His Glu Tyr Asn Thr Gln Phe Gln Glu Lys Ser Arg Glu Tyr Asp				
	450		455	460
Arg Leu Tyr Glu Glu Tyr Thr Arg Thr Ser Gln Glu Ile Gln Met Lys				
465		470		475
Arg Thr Ala Ile Glu Ala Phe Asn Glu Thr Ile Lys Ile Phe Glu Glu				
	485		490	495
Gln Cys Gln Thr Gln Glu Arg Tyr Ser Lys Glu Tyr Ile Glu Lys Phe				
	500		505	510
Lys Arg Glu Gly Asn Glu Lys Glu Ile Gln Arg Ile Met His Asn Tyr				
	515		520	525
Asp Lys Leu Lys Ser Arg Ile Ser Glu Ile Ile Asp Ser Arg Arg Arg				
	530		535	540
Leu Glu Glu Asp Leu Lys Lys Gln Ala Ala Glu Tyr Arg Glu Ile Asp				
545		550		555
Lys Arg Met Asn Ser Ile Lys Pro Asp Leu Ile Gln Leu Arg Lys Thr				
	565		570	575
Arg Asp Gln Tyr Leu Met Trp Leu Thr Gln Lys Gly Val Arg Gln Lys				
	580		585	590
Lys Leu Asn Glu Trp Leu Gly Asn Glu Asn Thr Glu Asp Gln Tyr Ser				
	595		600	605
Leu Val Glu Asp Asp Glu Asp Leu Pro His His Asp Glu Lys Thr Trp				
	610		615	620
Asn Val Gly Ser Ser Asn Arg Asn Lys Ala Glu Asn Leu Leu Arg Gly				
625		630		635
Lys Arg Asp Gly Thr Phe Leu Val Arg Glu Ser Ser Lys Gln Gly Cys				
	645		650	655
Tyr Ala Cys Ser Val Val Val Asp Gly Glu Val Lys His Cys Val Ile				
	660		665	670
Asn Lys Thr Ala Thr Gly Tyr Gly Phe Ala Glu Pro Tyr Asn Leu Tyr				

675	680	685
Ser Ser Leu Lys Glu Leu Val Leu His Tyr Gln His Thr Ser Leu Val		
690	695	700
Gln His Asn Asp Ser Leu Asn Val Thr Leu Ala Tyr Pro Val Tyr Ala		
705	710	715
Gln Gln Arg Arg Gln Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly		
725	730	735
Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly		
740	745	750
Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp		
755	760	765
Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys		
770	775	780
Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val		
785	790	795
Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe		
805	810	815
Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe		
820	825	830
Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly		
835	840	845
Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu		
850	855	860
Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His		
865	870	875
Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn		
885	890	895
Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp		
900	905	910
His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro		
915	920	925
Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn		
930	935	940
Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly		
945	950	955
Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys		
965	970	

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1788 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1785
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

ATG GGC AAC GCC GCC GCC GCC AAG AAG GGC AGC GAG CAG GAG AGC GTG
Met Gly Asn Ala Ala Ala Ala Lys Lys Gly Ser Glu Gln Glu Ser Val
1 5 10 15

31

AAA GAG TTC CTA GCC AAA GCC AAG GAA GAT TTC CTG AAA AAA TGG GAA Lys Glu Phe Leu Ala Lys Ala Lys Glu Asp Phe Leu Lys Lys Trp Glu 20 25 30	96
GAC CCC TCT CAG AAT ACA GCC CAG TTG GAT CAG TTT GAT AGA ATC AAG Asp Pro Ser Gln Asn Thr Ala Gln Leu Asp Gln Phe Asp Arg Ile Lys 35 40 45	144
ACC CTT GGC ACC GGC TCC TTT GGG CGA GTG ATG CTG GTG AAG CAC AAG Thr Leu Gly Thr Gly Ser Phe Gly Arg Val Met Leu Val Lys His Lys 50 55 60	192
GAG AGT GGG AAC CAC TAC GCC ATG AAG ATC TTA GAC AAG CAG AAG GTG Glu Ser Gly Asn His Tyr Ala Met Lys Ile Leu Asp Lys Gln Lys Val 65 70 75 80	240
GTG AAG CTA AAG CAG ATC GAG CAC ACT CTG AAT GAG AAG CGC ATC CTG Val Lys Leu Lys Gln Ile Glu His Thr Leu Asn Glu Lys Arg Ile Leu 85 90 95	288
CAG GCC GTC AAC TTC CCG TTC CTG GTC AAA CTT GAA TTC TCC TTC AAG Gln Ala Val Asn Phe Pro Phe Leu Val Lys Leu Glu Phe Ser Phe Lys 100 105 110	336
GAC AAC TCA AAC CTG TAC ATG GTC ATG GAG TAT GTA GCT GGT GGC GAG Asp Asn Ser Asn Leu Tyr Met Val Met Glu Tyr Val Ala Gly Gly Glu 115 120 125	384
ATG TTC TCC CAC CTA CGG CGG ATT GGA AGG TTC AGC GAG CCC CAT GCC Met Phe Ser His Leu Arg Arg Ile Gly Arg Phe Ser Glu Pro His Ala 130 135 140	432
CGT TTC TAC GCG GCG CAG ATC GTC CTG ACC TTT GAG TAT CTG CAC TCC Arg Phe Tyr Ala Ala Gln Ile Val Leu Thr Phe Glu Tyr Leu His Ser 145 150 155 160	480
CTG GAC CTC ATC TAC CGG GAC CTG AAG CCC GAG AAT CTT CTC ATC GAC Leu Asp Leu Ile Tyr Arg Asp Leu Lys Pro Glu Asn Leu Leu Ile Asp 165 170 175	528
CAG CAG GGC TAT ATT CAG GTG ACA GAC TTC GGT TTT GCC AAG CGT GTG Gln Gln Gly Tyr Ile Gln Val Thr Asp Phe Gly Phe Ala Lys Arg Val 180 185 190	576
AAA GGC CGT ACT TGG ACC TTG TGT GGG ACC CCT GAG TAC TTG GCC CCC Lys Gly Arg Thr Trp Thr Leu Cys Gly Thr Pro Glu Tyr Leu Ala Pro 195 200 205	624
GAG ATT ATC CTG AGC AAA GGC TAC AAC AAG GCT GTG GAC TGG TGG GCT Glu Ile Ile Leu Ser Lys Gly Tyr Asn Lys Ala Val Asp Trp Trp Ala 210 215 220	672
CTC GGA GTC CTC ATC TAC GAG ATG GCT GCT GGT TAC CCA CCC TTC TTC Leu Gly Val Leu Ile Tyr Glu Met Ala Ala Gly Tyr Pro Pro Phe Phe 225 230 235 240	720
GCT GAC CAG CCT ATC CAG ATC TAT GAG AAA ATC GTC TCT GGG AAG GTG	768

Ala Asp Gln Pro Ile Gln Ile Tyr Glu Lys Ile Val Ser Gly Lys Val	
245 250 255	
CGG TTC CCA TCC CAC TTC AGC TCT GAC TTG AAG GAC CTG CTG CGG AAC	816
Arg Phe Pro Ser His Phe Ser Ser Asp Leu Lys Asp Leu Leu Arg Asn	
260 265 270	
CTT CTG CAA GTG GAT CTA ACC AAG CGC TTT GGA AAC CTC AAG GAC GGG	864
Leu Leu Gln Val Asp Leu Thr Lys Arg Phe Gly Asn Leu Lys Asp Gly	
275 280 285	
GTC AAT GAC ATC AAG AAC CAC AAG TGG TTT GCC ACG ACT GAC TGG ATT	912
Val Asn Asp Ile Lys Asn His Lys Trp Phe Ala Thr Thr Asp Trp Ile	
290 295 300	
GCC ATC TAT CAG AGA AAG GTG GAA GCT CCC TTC ATA CCA AAG TTT AAA	960
Ala Ile Tyr Gln Arg Lys Val Glu Ala Pro Phe Ile Pro Lys Phe Lys	
305 310 315 320	
GGC CCT GGG GAC ACG AGT AAC TTT GAC GAC TAT GAG GAG GAA GAG ATC	1008
Gly Pro Gly Asp Thr Ser Asn Phe Asp Asp Tyr Glu Glu Glu Glu Ile	
325 330 335	
CGG GTC TCC ATC AAT GAG AAG TGT GGC AAG GAG TTT ACT GAG TTT GGG	1056
Arg Val Ser Ile Asn Glu Lys Cys Gly Lys Glu Phe Thr Glu Phe Gly	
340 345 350	
CGC GCC ATG AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT GTC CCA ATT	1104
Arg Ala Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile	
355 360 365	
CTT GTT GAA TTA GAT GGC GAT GTT AAT GGG CAA AAA TTC TCT GTT AGT	1152
Leu Val Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser	
370 375 380	
GGA GAG GGT GAA GGT GAT GCA ACA TAC GGA AAA CTT ACC CTT AAA TTT	1200
Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe	
385 390 395 400	
ATT TGC ACT ACT GGG AAG CTA CCT GTT CCA TGG CCA ACG CTT GTC ACT	1248
Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr	
405 410 415	
ACT CTC ACT TAT GGT GTT CAA TGC TTT TCT AGA TAC CCA GAT CAT ATG	1296
Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met	
420 425 430	
AAA CAG CAT GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT TAT GTA CAG	1344
Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln	
435 440 445	
GAA AGA ACT ATA TTT TAC AAA GAT GAC GGG AAC TAC AAG ACA CGT GCT	1392
Glu Arg Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala	
450 455 460	
GAA GTC AAG TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG TTA AAA	1440
Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys	
465 470 475 480	

GGT ATT GAT TTT AAA GAA GAT GGA AAC ATT CTT GGA CAC AAA ATG GAA Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu 485 490 495	1488
TAC AAT TAT AAC TCA CAT AAT GTA TAC ATC ATG GCA GAC AAA CCA AAG Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro Lys 500 505 510	1536
AAT GGC ATC AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT AAA GAT GGA Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly 515 520 525	1584
AGC GTT CAA TTA GCA GAC CAT TAT CAA CAA AAT ACT CCA ATT GGC GAT Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp 530 535 540	1632
GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACG CAA TCT GCC Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala 545 550 555 560	1680
CTT TCC AAA GAT CCC AAC GAA AAG AGA GAT CAC ATG ATC CTT CTT GAG Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu 565 570 575	1728
TTT GTA ACA GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA TAC AAA Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys 580 585 590	1776
CCT CAG GAG TAA Pro Gln Glu 595	1788

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 595 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Met Gly Asn Ala Ala Ala Lys Lys Gly Ser Glu Gln Glu Ser Val 1 5 10 15
Lys Glu Phe Leu Ala Lys Ala Lys Glu Asp Phe Leu Lys Lys Trp Glu 20 25 30
Asp Pro Ser Gln Asn Thr Ala Gln Leu Asp Gln Phe Asp Arg Ile Lys 35 40 45
Thr Leu Gly Thr Gly Ser Phe Gly Arg Val Met Leu Val Lys His Lys 50 55 60
Glu Ser Gly Asn His Tyr Ala Met Lys Ile Leu Asp Lys Gln Lys Val 65 70 75 80
Val Lys Leu Lys Gln Ile Glu His Thr Leu Asn Glu Lys Arg Ile Leu

				85						90					95
Gln	Ala	Val	Asn	Phe	Pro	Phe	Leu	Val	Lys	Leu	Glu	Phe	Ser	Phe	Lys
			100						105					110	
Asp	Asn	Ser	Asn	Leu	Tyr	Met	Val	Met	Glu	Tyr	Val	Ala	Gly	Gly	Glu
			115						120					125	
Met	Phe	Ser	His	Leu	Arg	Arg	Ile	Gly	Arg	Phe	Ser	Glu	Pro	His	Ala
			130						135					140	
Arg	Phe	Tyr	Ala	Ala	Gln	Ile	Val	Leu	Thr	Phe	Glu	Tyr	Leu	His	Ser
			145						150					155	160
Leu	Asp	Leu	Ile	Tyr	Arg	Asp	Leu	Lys	Pro	Glu	Asn	Leu	Leu	Ile	Asp
			165						170					175	
Gln	Gln	Gly	Tyr	Ile	Gln	Val	Thr	Asp	Phe	Gly	Phe	Ala	Lys	Arg	Val
			180						185					190	
Lys	Gly	Arg	Thr	Trp	Thr	Leu	Cys	Gly	Thr	Pro	Glu	Tyr	Leu	Ala	Pro
			195						200					205	
Glu	Ile	Ile	Leu	Ser	Lys	Gly	Tyr	Asn	Lys	Ala	Val	Asp	Trp	Trp	Ala
			210						215					220	
Leu	Gly	Val	Leu	Ile	Tyr	Glu	Met	Ala	Ala	Gly	Tyr	Pro	Pro	Phe	Phe
			225						230					235	240
Ala	Asp	Gln	Pro	Ile	Gln	Ile	Tyr	Glu	Lys	Ile	Val	Ser	Gly	Lys	Val
			245						250					255	
Arg	Phe	Pro	Ser	His	Phe	Ser	Ser	Asp	Leu	Lys	Asp	Leu	Leu	Arg	Asn
			260						265					270	
Leu	Leu	Gln	Val	Asp	Leu	Thr	Lys	Arg	Phe	Gly	Asn	Leu	Lys	Asp	Gly
			275						280					285	
Val	Asn	Asp	Ile	Lys	Asn	His	Lys	Trp	Phe	Ala	Thr	Thr	Asp	Trp	Ile
			290						295					300	
Ala	Ile	Tyr	Gln	Arg	Lys	Val	Glu	Ala	Pro	Phe	Ile	Pro	Lys	Phe	Lys
			305						310					315	320
Gly	Pro	Gly	Asp	Thr	Ser	Asn	Phe	Asp	Asp	Tyr	Glu	Glu	Glu	Glu	Ile
			325						330					335	
Arg	Val	Ser	Ile	Asn	Glu	Lys	Cys	Gly	Lys	Glu	Phe	Thr	Glu	Phe	Gly
			340						345					350	
Arg	Ala	Met	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile
			355						360					365	
Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	Gln	Lys	Phe	Ser	Val	Ser
			370						375					380	
Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe
			385						390					395	400
Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr
			405						410					415	
Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met
			420						425					430	
Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln
			435						440					445	
Glu	Arg	Thr	Ile	Phe	Tyr	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala
			450						455					460	
Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys
			465						470					475	480
Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Met	Glu
			485						490					495	
Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Pro	Lys
			500						505					510	
Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Lys	Asp	Gly
			515						520					525	
Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp
			530						535					540	
Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala

545 550 555 560
 Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu
 565 570 575
 Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
 580 585 590
 Pro Gln Glu
 595

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2181 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2178
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

ATG AGC GAC GTG GCT ATT GTG AAG GAG GGT TGG CTG CAC AAA CGA GGG	48
Met Ser Asp Val Ala Ile Val Lys Glu Gly Trp Leu His Lys Arg Gly	
1 5 10 15	
GAG TAC ATC AAG ACC TGG CGG CCA CGC TAC TTC CTC CTC AAG AAT GAT	96
Glu Tyr Ile Lys Thr Trp Arg Pro Arg Tyr Phe Leu Leu Lys Asn Asp	
20 25 30	
GGC ACC TTC ATT GGC TAC AAG GAG CGG CCG CAG GAT GTG GAC CAA CGT	144
Gly Thr Phe Ile Gly Tyr Lys Glu Arg Pro Gln Asp Val Asp Gln Arg	
35 40 45	
GAG GCT CCC CTC AAC AAC TTC TCT GTG GCG CAG TGC CAG CTG ATG AAG	192
Glu Ala Pro Leu Asn Asn Phe Ser Val Ala Gln Cys Gln Leu Met Lys	
50 55 60	
ACG GAG CGG CCC CGG CCC AAC ACC TTC ATC ATC CGC TGC CTG CAG TGG	240
Thr Glu Arg Pro Arg Pro Asn Thr Phe Ile Ile Arg Cys Leu Gln Trp	
65 70 75 80	
ACC ACT GTC ATC GAA CGC ACC TTC CAT GTG GAG ACT CCT GAG GAG CGG	288
Thr Thr Val Ile Glu Arg Thr Phe His Val Glu Thr Pro Glu Glu Arg	
85 90 95	
GAG GAG TGG ACA ACC GCC ATC CAG ACT GTG GCT GAC GGC CTC AAG AAG	336
Glu Glu Trp Thr Thr Ala Ile Gln Thr Val Ala Asp Gly Leu Lys Lys	
100 105 110	
CAG GAG GAG GAG GAG ATG GAC TTC CGG TCG GGC TCA CCC AGT GAC AAC	384
Gln Glu Glu Glu Glu Met Asp Phe Arg Ser Gly Ser Pro Ser Asp Asn	
115 120 125	
TCA GGG GCT GAA GAG ATG GAG GTG TCC CTG GCC AAG CCC AAG CAC CGC	432

Ser Gly Ala Glu Glu Met Glu Val Ser Leu Ala Lys Pro Lys His Arg	
130	135 140
GTG ACC ATG AAC GAG TTT GAG TAC CTG AAG CTG CTG GGC AAG GGC ACT	480
Val Thr Met Asn Glu Phe Glu Tyr Leu Lys Leu Leu Gly Lys Gly Thr	
145	150 155 160
TTC GGC AAG GTG ATC CTG GTG AAG GAG AAG GCC ACA GGC CGC TAC TAC	528
Phe Gly Lys Val Ile Leu Val Lys Glu Lys Ala Thr Gly Arg Tyr Tyr	
165	170 175
GCC ATG AAG ATC CTC AAG AAG GAA GTC ATC GTG GCC AAG GAC GAG GTG	576
Ala Met Lys Ile Leu Lys Lys Glu Val Ile Val Ala Lys Asp Glu Val	
180	185 190
GCC CAC ACA CTC ACC GAG AAC CGC GTC CTG CAG AAC TCC AGG CAC CCC	624
Ala His Thr Leu Thr Glu Asn Arg Val Leu Gln Asn Ser Arg His Pro	
195	200 205
TTC CTC ACA GCC CTG AAG TAC TCT TTC CAG ACC CAC GAC CGC CTC TGC	672
Phe Leu Thr Ala Leu Lys Tyr Ser Phe Gln Thr His Asp Arg Leu Cys	
210	215 220
TTT GTC ATG GAG TAC GCC AAC GGG GGC GAG CTG TTC TTC CAC CTG TCC	720
Phe Val Met Glu Tyr Ala Asn Gly Gly Glu Leu Phe Phe His Leu Ser	
225	230 235 240
CGG GAA CGT GTG TTC TCC GAG GAC CGG GCC CGC TTC TAT GGC GCT GAG	768
Arg Glu Arg Val Phe Ser Glu Asp Arg Ala Arg Phe Tyr Gly Ala Glu	
245	250 255
ATT GTG TCA GCC CTG GAC TAC CTG CAC TCG GAG AAG AAC GTG GTG TAC	816
Ile Val Ser Ala Leu Asp Tyr Leu His Ser Glu Lys Asn Val Val Tyr	
260	265 270
CGG GAC CTC AAG CTG GAG AAC CTC ATG CTG GAC AAG GAC GGG CAC ATT	864
Arg Asp Leu Lys Leu Glu Asn Leu Met Leu Asp Lys Asp Gly His Ile	
275	280 285
AAG ATC ACA GAC TTC GGG CTG TGC AAG GAG GGG ATC AAG GAC GGT GCC	912
Lys Ile Thr Asp Phe Gly Leu Cys Lys Glu Gly Ile Lys Asp Gly Ala	
290	295 300
ACC ATG AAG ACC TTT TGC GGC ACA CCT GAG TAC CTG GCC CCC GAG GTG	960
Thr Met Lys Thr Phe Cys Gly Thr Pro Glu Tyr Leu Ala Pro Glu Val	
305	310 315 320
CTG GAG GAC AAT GAC TAC GGC CGT GCA GTG GAC TGG TGG GGG CTG GGC	1008
Leu Glu Asp Asn Asp Tyr Gly Arg Ala Val Asp Trp Trp Gly Leu Gly	
325	330 335
GTG GTC ATG TAC GAG ATG ATG TGC GGT CGC CTG CCC TTC TAC AAC CAG	1056
Val Val Met Tyr Glu Met Met Cys Gly Arg Leu Pro Phe Tyr Asn Gln	
340	345 350
GAC CAT GAG AAG CTT TTT GAG CTC ATC CTC ATG GAG GAG ATC CGC TTC	1104
Asp His Glu Lys Leu Phe Glu Leu Ile Leu Met Glu Glu Ile Arg Phe	
355	360 365

CCG CGC ACG CTT GGT CCC GAG GCC AAG TCC TTG CTT TCA GGG CTG CTC	1152
Pro Arg Thr Leu Gly Pro Glu Ala Lys Ser Leu Leu Ser Gly Leu Leu	
370 375 380	
AAG AAG GAC CCC AAG CAG AGG CTT GGC GGG GGC TCC GAG GAC GCC AAG	1200
Lys Lys Asp Pro Lys Gln Arg Leu Gly Gly Ser Glu Asp Ala Lys	
385 390 395 400	
GAG ATC ATG CAG CAT CGC TTC TTT GCC GGT ATC GTG TGG CAG CAC GTG	1248
Glu Ile Met Gln His Arg Phe Phe Ala Gly Ile Val Trp Gln His Val	
405 410 415	
TAC GAG AAG AAG CTC AGC CCA CCC TTC AAG CCC CAG GTC ACG TCG GAG	1296
Tyr Glu Lys Lys Leu Ser Pro Pro Phe Lys Pro Gln Val Thr Ser Glu	
420 425 430	
ACT GAC ACC AGG TAT TTT GAT GAG GAG TTC ACG GCC CAG ATG ATC ACC	1344
Thr Asp Thr Arg Tyr Phe Asp Glu Glu Phe Thr Ala Gln Met Ile Thr	
435 440 445	
ATC ACA CCA CCT GAC CAA GAT GAC AGC ATG GAG TGT GTG GAC AGC GAG	1392
Ile Thr Pro Pro Asp Gln Asp Asp Ser Met Glu Cys Val Asp Ser Glu	
450 455 460	
CGC AGG CCC CAC TTC CCC CAG TTC TCC TAC TCG GCC AGC AGC ACG GCC	1440
Arg Arg Pro His Phe Pro Gln Phe Ser Tyr Ser Ala Ser Ser Thr Ala	
465 470 475 480	
TCG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC	1488
Ser Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe	
485 490 495	
ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC	1536
Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly	
500 505 510	
CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC	1584
His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly	
515 520 525	
AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC	1632
Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro	
530 535 540	
TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC	1680
Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser	
545 550 555 560	
CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG	1728
Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met	
565 570 575	
CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC	1776
Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly	
580 585 590	
AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG	1824

Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val	
595 600 605	
AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC	1872
Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile	
610 615 620	
CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC	1920
Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile	
625 630 635 640	
ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC	1968
Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg	
645 650 655	
CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG	2016
His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln	
660 665 670	
AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC	2064
Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr	
675 680 685	
CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT	2112
Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp	
690 695 700	
CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC	2160
His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly	
705 710 715 720	
ATG GAC GAG CTG TAC AAG TAA	2181
Met Asp Glu Leu Tyr Lys	
725	

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 726 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

Met Ser Asp Val Ala Ile Val Lys Glu Gly Trp Leu His Lys Arg Gly	
1 5 10 15	
Glu Tyr Ile Lys Thr Trp Arg Pro Arg Tyr Phe Leu Leu Lys Asn Asp	
20 25 30	
Gly Thr Phe Ile Gly Tyr Lys Glu Arg Pro Gln Asp Val Asp Gln Arg	
35 40 45	
Glu Ala Pro Leu Asn Asn Phe Ser Val Ala Gln Cys Gln Leu Met Lys	
50 55 60	
Thr Glu Arg Pro Arg Pro Asn Thr Phe Ile Ile Arg Cys Leu Gln Trp	

65					70					75				80	
Thr	Thr	Val	Ile	Glu	Arg	Thr	Phe	His	Val	Glu	Thr	Pro	Glu	Glu	Arg
				85						90				95	
Glu	Glu	Trp	Thr	Thr	Ala	Ile	Gln	Thr	Val	Ala	Asp	Gly	Leu	Lys	Lys
			100					105					110		
Gln	Glu	Glu	Glu	Glu	Met	Asp	Phe	Arg	Ser	Gly	Ser	Pro	Ser	Asp	Asn
			115				120					125			
Ser	Gly	Ala	Glu	Glu	Met	Glu	Val	Ser	Leu	Ala	Lys	Pro	Lys	His	Arg
	130					135					140				
Val	Thr	Met	Asn	Glu	Phe	Glu	Tyr	Leu	Lys	Leu	Gly	Lys	Gly	Thr	
145					150					155				160	
Phe	Gly	Lys	Val	Ile	Leu	Val	Lys	Glu	Lys	Ala	Thr	Gly	Arg	Tyr	Tyr
			165						170					175	
Ala	Met	Lys	Ile	Leu	Lys	Lys	Glu	Val	Ile	Val	Ala	Lys	Asp	Glu	Val
			180					185					190		
Ala	His	Thr	Leu	Thr	Glu	Asn	Arg	Val	Leu	Gln	Asn	Ser	Arg	His	Pro
	195					200						205			
Phe	Leu	Thr	Ala	Leu	Lys	Tyr	Ser	Phe	Gln	Thr	His	Asp	Arg	Leu	Cys
	210					215					220				
Phe	Val	Met	Glu	Tyr	Ala	Asn	Gly	Gly	Glu	Leu	Phe	Phe	His	Leu	Ser
225					230					235				240	
Arg	Glu	Arg	Val	Phe	Ser	Glu	Asp	Arg	Ala	Arg	Phe	Tyr	Gly	Ala	Glu
			245						250					255	
Ile	Val	Ser	Ala	Leu	Asp	Tyr	Leu	His	Ser	Glu	Lys	Asn	Val	Val	Tyr
			260					265					270		
Arg	Asp	Leu	Lys	Leu	Glu	Asn	Leu	Met	Leu	Asp	Lys	Asp	Gly	His	Ile
	275					280						285			
Lys	Ile	Thr	Asp	Phe	Gly	Leu	Cys	Lys	Glu	Gly	Ile	Lys	Asp	Gly	Ala
	290					295					300				
Thr	Met	Lys	Thr	Phe	Cys	Gly	Thr	Pro	Glu	Tyr	Leu	Ala	Pro	Glu	Val
305					310					315				320	
Leu	Glu	Asp	Asn	Asp	Tyr	Gly	Arg	Ala	Val	Asp	Trp	Trp	Gly	Leu	Gly
			325						330					335	
Val	Val	Met	Tyr	Glu	Met	Met	Cys	Gly	Arg	Leu	Pro	Phe	Tyr	Asn	Gln
		340					345						350		
Asp	His	Glu	Lys	Leu	Phe	Glu	Leu	Ile	Leu	Met	Glu	Glu	Ile	Arg	Phe
		355				360						365			
Pro	Arg	Thr	Leu	Gly	Pro	Glu	Ala	Lys	Ser	Leu	Leu	Ser	Gly	Leu	Leu
	370					375					380				
Lys	Lys	Asp	Pro	Lys	Gln	Arg	Leu	Gly	Gly	Gly	Ser	Glu	Asp	Ala	Lys
385					390					395				400	
Glu	Ile	Met	Gln	His	Arg	Phe	Phe	Ala	Gly	Ile	Val	Trp	Gln	His	Val
			405						410					415	
Tyr	Glu	Lys	Lys	Leu	Ser	Pro	Pro	Phe	Lys	Pro	Gln	Val	Thr	Ser	Glu
		420						425					430		
Thr	Asp	Thr	Arg	Tyr	Phe	Asp	Glu	Glu	Phe	Thr	Ala	Gln	Met	Ile	Thr
		435					440						445		
Ile	Thr	Pro	Pro	Asp	Gln	Asp	Asp	Ser	Met	Glu	Cys	Val	Asp	Ser	Glu
	450					455					460				
Arg	Arg	Pro	His	Phe	Pro	Gln	Phe	Ser	Tyr	Ser	Ala	Ser	Ser	Thr	Ala
465					470					475				480	
Ser	Asp	Pro	Pro	Val	Ala	Thr	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe
			485						490					495	
Thr	Gly	Val	Val	Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly
		500						505					510		
His	Lys	Phe	Ser	Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly
	515						520					525			
Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro

530		535		540											
Trp	Pro	Thr	Leu	Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser
545					550					555					560
Arg	Tyr	Pro	Asp	His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met
					565					570					575
Pro	Glu	Gly	Tyr	Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly
					580					585					590
Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val
					595					600					605
Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile
					610					615					620
Leu	Gly	His	Lys	Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile
					625					630					640
Met	Ala	Asp	Lys	Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg
					645					650					655
His	Asn	Ile	Glu	Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln
					660					665					670
Asn	Thr	Pro	Ile	Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr
					675					680					685
Leu	Ser	Thr	Gln	Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp
					690					695					700
His	Met	Val	Leu	Leu	Glu	Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly
					705					710					715
Met	Asp	Glu	Leu	Tyr	Lys										720
					725										

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2751 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2748
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

ATG	GCT	GAC	GTT	TAC	CCG	GCC	AAC	GAC	TCC	ACG	GCG	TCT	CAG	GAC	GTG	48
Met	Ala	Asp	Val	Tyr	Pro	Ala	Asn	Asp	Ser	Thr	Ala	Ser	Gln	Asp	Val	
1			5						10					15		
GCC	AAC	CGC	TTC	GCC	CGC	AAA	GGG	GCG	CTG	AGG	CAG	AAG	AAC	GTG	CAT	96
Ala	Asn	Arg	Phe	Ala	Arg	Lys	Gly	Ala	Leu	Arg	Gln	Lys	Asn	Val	His	
			20					25					30			
GAG	GTG	AAA	GAC	CAC	AAA	TTC	ATC	GCC	CGC	TTC	TTC	AAG	CAA	CCC	ACC	144
Glu	Val	Lys	Asp	His	Lys	Phe	Ile	Ala	Arg	Phe	Phe	Lys	Gln	Pro	Thr	
			35					40					45			
TTC	TGC	AGC	CAC	TGC	ACC	GAC	TTC	ATC	TGG	GGG	TTT	GGG	AAA	CAA	GGC	192
Phe	Cys	Ser	His	Cys	Thr	Asp	Phe	Ile	Trp	Gly	Phe	Gly	Lys	Gln	Gly	
			50					55							60	

TTC CAG TGC CAA GTT TGC TGT TTT GTG GTT CAT AAG AGG TGC CAT GAG Phe Gln Cys Gln Val Cys Cys Phe Val Val His Lys Arg Cys His Glu 65 70 75 80	240
TTC GTT ACG TTC TCT TGT CCG GGT GCG GAT AAG GGA CCT GAC ACT GAC Phe Val Thr Phe Ser Cys Pro Gly Ala Asp Lys Gly Pro Asp Thr Asp 85 90 95	288
GAC CCC AGG AGC AAG CAC AAG TTC AAA ATC CAC ACA TAC GGA AGC CCT Asp Pro Arg Ser Lys His Lys Phe Lys Ile His Thr Tyr Gly Ser Pro 100 105 110	336
ACC TTC TGT GAT CAC TGT GGG TCC CTG CTC TAT GGA CTT ATC CAC CAA Thr Phe Cys Asp His Cys Gly Ser Leu Leu Tyr Gly Leu Ile His Gln 115 120 125	384
GGG ATG AAA TGT GAC ACC TGC GAC ATG AAT GTT CAC AAC CAG TGT GTG Gly Met Lys Cys Asp Thr Cys Asp Met Asn Val His Asn Gln Cys Val 130 135 140	432
ATC AAT GAC CCT AGC CTC TGC GGA ATG GAT CAC ACA GAG AAG AGG GGG Ile Asn Asp Pro Ser Leu Cys Gly Met Asp His Thr Glu Lys Arg Gly 145 150 155 160	480
CGG ATT TAT CTG AAG GCT GAG GTC ACT GAT GAA AAG CTC CAC GTC ACG Arg Ile Tyr Leu Lys Ala Glu Val Thr Asp Glu Lys Leu His Val Thr 165 170 175	528
GTA CGA GAT GCA AAA AAT CTA ATC CCT ATG GAT CCA AAT GGG CTT TCG Val Arg Asp Ala Lys Asn Leu Ile Pro Met Asp Pro Asn Gly Leu Ser 180 185 190	576
GAT CCT TAT GTG AAG CTG AAA CTA ATC CCT GAC CCC AAG AAT GAG AGC Asp Pro Tyr Val Lys Leu Lys Leu Ile Pro Asp Pro Lys Asn Glu Ser 195 200 205	624
AAA CAG AAA ACC AAA ACC ATC CGC TCC AAC CTG AAT CCT CAG TGG AAT Lys Gln Lys Thr Lys Thr Ile Arg Ser Asn Leu Asn Pro Gln Trp Asn 210 215 220	672
GAG TCC TTC ACG TTC AAA TTA AAA CCT TCA GAC AAA GAC CGG CGA CTG Glu Ser Phe Thr Phe Lys Leu Lys Pro Ser Asp Lys Asp Arg Arg Leu 225 230 235 240	720
TCT GTA GAA ATC TGG GAC TGG GAT CGG ACG ACT CGG AAT GAC TTC ATG Ser Val Glu Ile Trp Asp Trp Asp Arg Thr Thr Arg Asn Asp Phe Met 245 250 255	768
GGA TCC CTT TCC TTT GGT GTC TCA GAG CTA ATG AAG ATG CCG GCC AGT Gly Ser Leu Ser Phe Gly Val Ser Glu Leu Met Lys Met Pro Ala Ser 260 265 270	816
GGA TGG TAT AAA GCT CAC AAC CAA GAA GAG GGC GAA TAT TAC AAC GTG Gly Trp Tyr Lys Ala His Asn Gln Glu Glu Gly Glu Tyr Tyr Asn Val 275 280 285	864
CCC ATT CCA GAA GGA GAT GAA GAA GGC AAC ATG GAA CTC AGG CAG AAG	912

Pro Ile Pro Glu Gly Asp Glu Glu Gly Asn Met Glu Leu Arg Gln Lys	
290 295 300	
TTT GAG AAA GCC AAG CTA GGT CCT GTT GGT AAC AAA GTC ATC AGC CCT	960
Phe Glu Lys Ala Lys Leu Gly Pro Val Gly Asn Lys Val Ile Ser Pro	
305 310 315 320	
TCA GAA GAC AGA AAG CAA CCA TCC AAC AAC CTG GAC AGA GTG AAA CTC	1008
Ser Glu Asp Arg Lys Gln Pro Ser Asn Asn Leu Asp Arg Val Lys Leu	
325 330 335	
ACA GAC TTC AAC TTC CTC ATG GTG CTG GGG AAG GGG AGT TTT GGG AAG	1056
Thr Asp Phe Asn Phe Leu Met Val Leu Gly Lys Gly Ser Phe Gly Lys	
340 345 350	
GTG ATG CTT GCT GAC AGG AAG GGA ACG GAG GAA CTG TAC GCC ATC AAG	1104
Val Met Leu Ala Asp Arg Lys Gly Thr Glu Glu Leu Tyr Ala Ile Lys	
355 360 365	
ATC CTG AAG AAG GAC GTG GTG ATC CAG GAC GAC GTG GAG TGC ACC	1152
Ile Leu Lys Lys Asp Val Val Ile Gln Asp Asp Asp Val Glu Cys Thr	
370 375 380	
ATG GTG GAG AAG CGC GTG CTG GCC CTG CTG GAC AAG CCG CCA TTT CTG	1200
Met Val Glu Lys Arg Val Leu Ala Leu Leu Asp Lys Pro Pro Phe Leu	
385 390 395 400	
ACA CAG CTG CAC TCC TGC TTC CAG ACA GTG GAC CGG CTG TAC TTC GTC	1248
Thr Gln Leu His Ser Cys Phe Gln Thr Val Asp Arg Leu Tyr Phe Val	
405 410 415	
ATG GAA TAC GTC AAC GGC GGG GAT CTT ATG TAC CAC ATT CAG CAA GTC	1296
Met Glu Tyr Val Asn Gly Gly Asp Leu Met Tyr His Ile Gln Gln Val	
420 425 430	
GGG AAA TTT AAG GAG CCA CAA GCA GTA TTC TAC GCA GCC GAG ATC TCC	1344
Gly Lys Phe Lys Glu Pro Gln Ala Val Phe Tyr Ala Ala Glu Ile Ser	
435 440 445	
ATC GGA CTG TTC TTC CTT CAT AAA AGA GGG ATC ATT TAC AGG GAT CTG	1392
Ile Gly Leu Phe Phe Leu His Lys Arg Gly Ile Ile Tyr Arg Asp Leu	
450 455 460	
AAG CTG AAC AAT GTC ATG CTG AAC TCA GAA GGG CAC ATC AAA ATC GCC	1440
Lys Leu Asn Asn Val Met Leu Asn Ser Glu Gly His Ile Lys Ile Ala	
465 470 475 480	
GAC TTC GGG ATG TGC AAG GAA CAC ATG ATG GAT GGA GTC ACG ACC AGG	1488
Asp Phe Gly Met Cys Lys Glu His Met Met Asp Gly Val Thr Thr Arg	
485 490 495	
ACC TTC TGC GGA ACT CCG GAC TAC ATT GCC CCA GAG ATA ATC GCT TAC	1536
Thr Phe Cys Gly Thr Pro Asp Tyr Ile Ala Pro Glu Ile Ile Ala Tyr	
500 505 510	
CAG CCG TAC GGG AAG TCT GTA GAT TGG TGG GCG TAC GGT GTG CTG CTG	1584
Gln Pro Tyr Gly Lys Ser Val Asp Trp Trp Ala Tyr Gly Val Leu Leu	
515 520 525	

TAC GAG ATG CTA GCC GGG CAG CCT CCG TTT GAT GGT GAA GAT GAA GAT	1632
Tyr Glu Met Leu Ala Gly Gln Pro Pro Phe Asp Gly Glu Asp Glu Asp	
530 535 540	
GAA CTG TTT CAG TCT ATA ATG GAG CAC AAC GTG TCC TAC CCC AAA TCC	1680
Glu Leu Phe Gln Ser Ile Met Glu His Asn Val Ser Tyr Pro Lys Ser	
545 550 555 560	
TTG TCC AAG GAA GCC GTC TCC ATC TGC AAA GGA CTT ATG ACC AAA CAG	1728
Leu Ser Lys Glu Ala Val Ser Ile Cys Lys Gly Leu Met Thr Lys Gln	
565 570 575	
CCT GCC AAG CGA CTG GGC TGC GGG CCC GAG GGA GAG AGG GAT GTC AGA	1776
Pro Ala Lys Arg Leu Gly Cys Gly Pro Glu Gly Glu Arg Asp Val Arg	
580 585 590	
GAG CAT GCC TTC TTC AGG AGG ATC GAC TGG GAG AAA CTG GAG AAC AGG	1824
Glu His Ala Phe Phe Arg Arg Ile Asp Trp Glu Lys Leu Glu Asn Arg	
595 600 605	
GAG ATC CAA CCA CCA TTC AAG CCC AAA GTG TGT GGC AAA GGA GCA GAA	1872
Glu Ile Gln Pro Pro Phe Lys Pro Lys Val Cys Gly Lys Gly Ala Glu	
610 615 620	
AAC TTT GAC AAG TTC TTC ACG CGA GGA CAG CCT GTC TTA ACA CCA CCA	1920
Asn Phe Asp Lys Phe Phe Thr Arg Gly Gln Pro Val Leu Thr Pro Pro	
625 630 635 640	
GAT CAG CTG GTC ATT GCT AAC ATA GAC CAA TCT GAT TTT GAA GGG TTC	1968
Asp Gln Leu Val Ile Ala Asn Ile Asp Gln Ser Asp Phe Glu Gly Phe	
645 650 655	
TCG TAT GTC AAC CCC CAG TTT GTG CAC CCA ATC TTG CAA AGT GCA GTA	2016
Ser Tyr Val Asn Pro Gln Phe Val His Pro Ile Leu Gln Ser Ala Val	
660 665 670	
GGG CGC GCC ATG AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT GTC CCA	2064
Gly Arg Ala Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro	
675 680 685	
ATT CTT GTT GAA TTA GAT GGC GAT GTT AAT GGG CAA AAA TTC TCT GTT	2112
Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val	
690 695 700	
AGT GGA GAG GGT GAA GGT GAT GCA ACA TAC GGA AAA CTT ACC CTT AAA	2160
Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys	
705 710 715 720	
TTT ATT TGC ACT ACT GGG AAG CTA CCT GTT CCA TGG CCA ACG CTT GTC	2208
Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val	
725 730 735	
ACT ACT CTC ACT TAT GGT GTT CAA TGC TTT TCT AGA TAC CCA GAT CAT	2256
Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His	
740 745 750	
ATG AAA CAG CAT GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT TAT GTA	2304

Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val	
755 760 765	
CAG GAA AGA ACT ATA TTT TAC AAA GAT GAC GGG AAC TAC AAG ACA CGT	2352
Gln Glu Arg Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg	
770 775 780	
GCT GAA GTC AAG TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG TTA	2400
Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu	
785 790 795 800	
AAA GGT ATT GAT TTT AAA GAA GAT GGA AAC ATT CTT GGA CAC AAA ATG	2448
Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met	
805 810 815	
GAA TAC AAT TAT AAC TCA CAT AAT GTA TAC ATC ATG GCA GAC AAA CCA	2496
Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Pro	
820 825 830	
AAG AAT GGC ATC AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT AAA GAT	2544
Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp	
835 840 845	
GGA AGC GTT CAA TTA GCA GAC CAT TAT CAA CAA AAT ACT CCA ATT GGC	2592
Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly	
850 855 860	
GAT GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACG CAA TCT	2640
Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser	
865 870 875 880	
GCC CTT TCC AAA GAT CCC AAC GAA AAG AGA GAT CAC ATG ATC CTT CTT	2688
Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu	
885 890 895	
GAG TTT GTA ACA GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA TAC	2736
Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr	
900 905 910	
AAA CCT CAG GAG TAA	2751
Lys Pro Gln Glu	
915	

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 916 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Met Ala Asp Val Tyr Pro Ala Asn Asp Ser Thr Ala Ser Gln Asp Val

1	5	10	15
Ala Asn Arg Phe Ala Arg Lys Gly Ala Leu Arg Gln Lys Asn Val His			
20	25	30	
Glu Val Lys Asp His Lys Phe Ile Ala Arg Phe Phe Lys Gln Pro Thr			
35	40	45	
Phe Cys Ser His Cys Thr Asp Phe Ile Trp Gly Phe Gly Lys Gln Gly			
50	55	60	
Phe Gln Cys Gln Val Cys Cys Phe Val Val His Lys Arg Cys His Glu			
65	70	75	80
Phe Val Thr Phe Ser Cys Pro Gly Ala Asp Lys Gly Pro Asp Thr Asp			
85	90	95	
Asp Pro Arg Ser Lys His Lys Phe Lys Ile His Thr Tyr Gly Ser Pro			
100	105	110	
Thr Phe Cys Asp His Cys Gly Ser Leu Leu Tyr Gly Leu Ile His Gln			
115	120	125	
Gly Met Lys Cys Asp Thr Cys Asp Met Asn Val His Asn Gln Cys Val			
130	135	140	
Ile Asn Asp Pro Ser Leu Cys Gly Met Asp His Thr Glu Lys Arg Gly			
145	150	155	160
Arg Ile Tyr Leu Lys Ala Glu Val Thr Asp Glu Lys Leu His Val Thr			
165	170	175	
Val Arg Asp Ala Lys Asn Leu Ile Pro Met Asp Pro Asn Gly Leu Ser			
180	185	190	
Asp Pro Tyr Val Lys Leu Lys Leu Ile Pro Asp Pro Lys Asn Glu Ser			
195	200	205	
Lys Gln Lys Thr Lys Thr Ile Arg Ser Asn Leu Asn Pro Gln Trp Asn			
210	215	220	
Glu Ser Phe Thr Phe Lys Leu Lys Pro Ser Asp Lys Asp Arg Arg Leu			
225	230	235	240
Ser Val Glu Ile Trp Asp Trp Asp Arg Thr Thr Arg Asn Asp Phe Met			
245	250	255	
Gly Ser Leu Ser Phe Gly Val Ser Glu Leu Met Lys Met Pro Ala Ser			
260	265	270	
Gly Trp Tyr Lys Ala His Asn Gln Glu Glu Gly Glu Tyr Tyr Asn Val			
275	280	285	
Pro Ile Pro Glu Gly Asp Glu Glu Gly Asn Met Glu Leu Arg Gln Lys			
290	295	300	
Phe Glu Lys Ala Lys Leu Gly Pro Val Gly Asn Lys Val Ile Ser Pro			
305	310	315	320
Ser Glu Asp Arg Lys Gln Pro Ser Asn Asn Leu Asp Arg Val Lys Leu			
325	330	335	
Thr Asp Phe Asn Phe Leu Met Val Leu Gly Lys Gly Ser Phe Gly Lys			
340	345	350	
Val Met Leu Ala Asp Arg Lys Gly Thr Glu Glu Leu Tyr Ala Ile Lys			
355	360	365	
Ile Leu Lys Lys Asp Val Val Ile Gln Asp Asp Asp Val Glu Cys Thr			
370	375	380	
Met Val Glu Lys Arg Val Leu Ala Leu Leu Asp Lys Pro Pro Phe Leu			
385	390	395	400
Thr Gln Leu His Ser Cys Phe Gln Thr Val Asp Arg Leu Tyr Phe Val			
405	410	415	
Met Glu Tyr Val Asn Gly Gly Asp Leu Met Tyr His Ile Gln Gln Val			
420	425	430	
Gly Lys Phe Lys Glu Pro Gln Ala Val Phe Tyr Ala Ala Glu Ile Ser			
435	440	445	
Ile Gly Leu Phe Phe Leu His Lys Arg Gly Ile Ile Tyr Arg Asp Leu			
450	455	460	
Lys Leu Asn Asn Val Met Leu Asn Ser Glu Gly His Ile Lys Ile Ala			

465	470								475								480	
Asp	Phe	Gly	Met	Cys	Lys	Glu	His	Met	Met	Asp	Gly	Val	Thr	Thr	Arg			
				485					490						495			
Thr	Phe	Cys	Gly	Thr	Pro	Asp	Tyr	Ile	Ala	Pro	Glu	Ile	Ile	Ala	Tyr			
			500					505					510					
Gln	Pro	Tyr	Gly	Lys	Ser	Val	Asp	Trp	Trp	Ala	Tyr	Gly	Val	Leu	Leu			
		515					520					525						
Tyr	Glu	Met	Leu	Ala	Gly	Gln	Pro	Pro	Phe	Asp	Gly	Glu	Asp	Glu	Asp			
	530					535					540							
Glu	Leu	Phe	Gln	Ser	Ile	Met	Glu	His	Asn	Val	Ser	Tyr	Pro	Lys	Ser			
545					550					555					560			
Leu	Ser	Lys	Glu	Ala	Val	Ser	Ile	Cys	Lys	Gly	Leu	Met	Thr	Lys	Gln			
				565					570					575				
Pro	Ala	Lys	Arg	Leu	Gly	Cys	Gly	Pro	Glu	Gly	Glu	Arg	Asp	Val	Arg			
			580					585					590					
Glu	His	Ala	Phe	Phe	Arg	Arg	Ile	Asp	Trp	Glu	Lys	Leu	Glu	Asn	Arg			
	595						600					605						
Glu	Ile	Gln	Pro	Pro	Phe	Lys	Pro	Lys	Val	Cys	Gly	Lys	Gly	Ala	Glu			
	610					615					620							
Asn	Phe	Asp	Lys	Phe	Phe	Thr	Arg	Gly	Gln	Pro	Val	Leu	Thr	Pro	Pro			
625					630					635					640			
Asp	Gln	Leu	Val	Ile	Ala	Asn	Ile	Asp	Gln	Ser	Asp	Phe	Glu	Gly	Phe			
				645					650					655				
Ser	Tyr	Val	Asn	Pro	Gln	Phe	Val	His	Pro	Ile	Leu	Gln	Ser	Ala	Val			
			660					665					670					
Gly	Arg	Ala	Met	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro			
	675					680						685						
Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	Gln	Lys	Phe	Ser	Val			
	690					695					700							
Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys			
705					710					715					720			
Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val			
				725					730					735				
Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His			
			740					745					750					
Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val			
	755						760					765						
Gln	Glu	Arg	Thr	Ile	Phe	Tyr	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg			
	770					775					780							
Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu			
785					790					795					800			
Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Met			
				805					810									

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2157 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2154
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

ATG TCG TCC ATC TTG CCA TTC ACG CCG CCA GTT GTG AAG AGA CTG CTG	48
Met Ser Ser Ile Leu Pro Phe Thr Pro Pro Val Val Lys Arg Leu Leu	
1 5 10 15	
GGA TGG AAG AAG TCA GCT GGT GGG TCT GGA GGA GCA GGC GGA GGA GAG	96
Gly Trp Lys Lys Ser Ala Gly Gly Ser Gly Gly Ala Gly Gly Gly Glu	
20 25 30	
CAG AAT GGG CAG GAA GAA AAG TGG TGT GAG AAA GCA GTG AAA AGT CTG	144
Gln Asn Gly Gln Glu Glu Lys Trp Cys Glu Lys Ala Val Lys Ser Leu	
35 40 45	
GTG AAG AAG CTA AAG AAA ACA GGA CGA TTA GAT GAG CTT GAG AAA GCC	192
Val Lys Lys Leu Lys Lys Thr Gly Arg Leu Asp Glu Leu Glu Lys Ala	
50 55 60	
ATC ACC ACT CAA AAC TGT AAT ACT AAA TGT GTT ACC ATA CCA AGC ACT	240
Ile Thr Thr Gln Asn Cys Asn Thr Lys Cys Val Thr Ile Pro Ser Thr	
65 70 75 80	
TGC TCT GAA ATT TGG GGA CTG AGT ACA CCA AAT ACG ATA GAT CAG TGG	288
Cys Ser Glu Ile Trp Gly Leu Ser Thr Pro Asn Thr Ile Asp Gln Trp	
85 90 95	
GAT ACA ACA GGC CTT TAC AGC TTC TCT GAA CAA ACC AGG TCT CTT GAT	336
Asp Thr Thr Gly Leu Tyr Ser Phe Ser Glu Gln Thr Arg Ser Leu Asp	
100 105 110	
GGT CGT CTC CAG GTA TCC CAT CGA AAA GGA TTG CCA CAT GTT ATA TAT	384
Gly Arg Leu Gln Val Ser His Arg Lys Gly Leu Pro His Val Ile Tyr	
115 120 125	
TGC CGA TTA TGG CGC TGG CCT GAT CTT CAC AGT CAT CAT GAA CTC AAG	432
Cys Arg Leu Trp Arg Trp Pro Asp Leu His Ser His His Glu Leu Lys	
130 135 140	
GCA ATT GAA AAC TGC GAA TAT GCT TTT AAT CTT AAA AAG GAT GAA GTA	480
Ala Ile Glu Asn Cys Glu Tyr Ala Phe Asn Leu Lys Lys Asp Glu Val	
145 150 155 160	
TGT GTA AAC CCT TAC CAC TAT CAG AGA GTT GAG ACA CCA GTT TTG CCT	528

Cys Val Asn Pro Tyr His Tyr Gln Arg Val Glu Thr Pro Val Leu Pro	
165 170 175	
CCA GTA TTA GTG CCC CGA CAC ACC GAG ATC CTA ACA GAA CTT CCG CCT	576
Pro Val Leu Val Pro Arg His Thr Glu Ile Leu Thr Glu Leu Pro Pro	
180 185 190	
CTG GAT GAC TAT ACT CAC TCC ATT CCA GAA AAC ACT AAC TTC CCA GCA	624
Leu Asp Asp Tyr Thr His Ser Ile Pro Glu Asn Thr Asn Phe Pro Ala	
195 200 205	
GGA ATT GAG CCA CAG AGT AAT TAT ATT CCA GAA ACG CCA CCT CCT GGA	672
Gly Ile Glu Pro Gln Ser Asn Tyr Ile Pro Glu Thr Pro Pro Pro Gly	
210 215 220	
TAT ATC AGT GAA GAT GGA GAA ACA AGT GAC CAA CAG TTG AAT CAA AGT	720
Tyr Ile Ser Glu Asp Gly Glu Thr Ser Asp Gln Gln Leu Asn Gln Ser	
225 230 235 240	
ATG GAC ACA GGC TCT CCA GCA GAA CTA TCT CCT ACT ACT CTT TCC CCT	768
Met Asp Thr Gly Ser Pro Ala Glu Leu Ser Pro Thr Thr Leu Ser Pro	
245 250 255	
GTT AAT CAT AGC TTG GAT TTA CAG CCA GTT ACT TAC TCA GAA CCT GCA	816
Val Asn His Ser Leu Asp Leu Gln Pro Val Thr Tyr Ser Glu Pro Ala	
260 265 270	
TTT TGG TGT TCA ATA GCA TAT TAT GAA TTA AAT CAG AGG GTT GGA GAA	864
Phe Trp Cys Ser Ile Ala Tyr Tyr Glu Leu Asn Gln Arg Val Gly Glu	
275 280 285	
ACC TTC CAT GCA TCA CAG CCC TCA CTC ACT GTA GAT GGC TTT ACA GAC	912
Thr Phe His Ala Ser Gln Pro Ser Leu Thr Val Asp Gly Phe Thr Asp	
290 295 300	
CCA TCA AAT TCA GAG AGG TTC TGC TTA GGT TTA CTC TCC AAT GTT AAC	960
Pro Ser Asn Ser Glu Arg Phe Cys Leu Gly Leu Leu Ser Asn Val Asn	
305 310 315 320	
CGA AAT GCC ACG GTA GAA ATG ACA AGA AGG CAT ATA GGA AGA GGA GTG	1008
Arg Asn Ala Thr Val Glu Met Thr Arg Arg His Ile Gly Arg Gly Val	
325 330 335	
CGC TTA TAC TAC ATA GGT GGG GAA GTT TTT GCT GAG TGC CTA ACT GAT	1056
Arg Leu Tyr Tyr Ile Gly Gly Glu Val Phe Ala Glu Cys Leu Ser Asp	
340 345 350	
AGT GCA ATC TTT GTG CAG AGC CCC AAT TGT AAT CAG AGA TAT GGC TGG	1104
Ser Ala Ile Phe Val Gln Ser Pro Asn Cys Asn Gln Arg Tyr Gly Trp	
355 360 365	
CAC CCT GCA ACA GTG TGT AAA ATT CCA CCA GGC TGT AAT CTG AAG ATC	1152
His Pro Ala Thr Val Cys Lys Ile Pro Pro Gly Cys Asn Leu Lys Ile	
370 375 380	
TTC AAC AAC CAG GAA TTT GCT GCT CTT CTG GCT CAG TCT GTT AAT CAG	1200
Phe Asn Asn Gln Glu Phe Ala Ala Leu Leu Ala Gln Ser Val Asn Gln	
385 390 395 400	

GGT TTT GAA GCC GTC TAT CAG CTA ACT AGA ATG TGC ACC ATA AGA ATG Gly Phe Glu Ala Val Tyr Gln Leu Thr Arg Met Cys Thr Ile Arg Met 405 410 415	1248
AGT TTT GTG AAA GGG TGG GGA GCA GAA TAC CGA AGG CAG ACG GTA ACA Ser Phe Val Lys Gly Trp Gly Ala Glu Tyr Arg Arg Gln Thr Val Thr 420 425 430	1296
AGT ACT CCT TGC TGG ATT GAA CTT CAT CTG AAT GGA CCT CTA CAG TGG Ser Thr Pro Cys Trp Ile Glu Leu His Leu Asn Gly Pro Leu Gln Trp 435 440 445	1344
TTG GAC AAA GTA TTA ACT CAG ATG GGA TCC CCT TCA GTG CGT TGC TCA Leu Asp Lys Val Leu Thr Gln Met Gly Ser Pro Ser Val Arg Cys Ser 450 455 460	1392
AGC ATG TCA TGG GTA CCG CGG GCC CGG GAT CCA CCG GTC GCC ACC ATG Ser Met Ser Trp Val Pro Arg Ala Arg Asp Pro Pro Val Ala Thr Met 465 470 475 480	1440
GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val 485 490 495	1488
GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu 500 505 510	1536
GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys 515 520 525	1584
ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu 530 535 540	1632
ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln 545 550 555 560	1680
CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg 565 570 575	1728
ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val 580 585 590	1776
AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile 595 600 605	1824
GAC TTC AAG GAG GAC GGC AAC ATC CTG GGC CAC AAG CTG GAG TAC AAC Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn 610 615 620	1872
TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC	1920

Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	Gly	
625					630					635					640	
ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	GTG	1968
Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	Val	
				645					650					655		
CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	CCC	2016
Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	Pro	
				660				665					670			
GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	TCC	GCC	CTG	AGC	2064
Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	Ser	
				675				680					685			
AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	CTG	GAG	TTC	GTG	2112
Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	Val	
				690				695					700			
ACC	GCC	GCC	GGG	ATC	ACT	CTC	GGC	ATG	GAC	GAG	CTG	TAC	AAG	TAA		2157
Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys			
705						710							715			

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 718 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Met	Ser	Ser	Ile	Leu	Pro	Phe	Thr	Pro	Pro	Val	Val	Lys	Arg	Leu	Leu	
1				5				10						15		
Gly	Trp	Lys	Lys	Ser	Ala	Gly	Gly	Ser	Gly	Gly	Ala	Gly	Gly	Gly	Glu	
			20					25						30		
Gln	Asn	Gly	Gln	Glu	Glu	Lys	Trp	Cys	Glu	Lys	Ala	Val	Lys	Ser	Leu	
			35				40					45				
Val	Lys	Lys	Leu	Lys	Lys	Thr	Gly	Arg	Leu	Asp	Glu	Leu	Glu	Lys	Ala	
			50			55					60					
Ile	Thr	Thr	Gln	Asn	Cys	Asn	Thr	Lys	Cys	Val	Thr	Ile	Pro	Ser	Thr	
65				70					75					80		
Cys	Ser	Glu	Ile	Trp	Gly	Leu	Ser	Thr	Pro	Asn	Thr	Ile	Asp	Gln	Trp	
			85					90						95		
Asp	Thr	Thr	Gly	Leu	Tyr	Ser	Phe	Ser	Glu	Gln	Thr	Arg	Ser	Leu	Asp	
			100					105					110			
Gly	Arg	Leu	Gln	Val	Ser	His	Arg	Lys	Gly	Leu	Pro	His	Val	Ile	Tyr	
			115				120					125				
Cys	Arg	Leu	Trp	Arg	Trp	Pro	Asp	Leu	His	Ser	His	His	Glu	Leu	Lys	
			130			135				140						
Ala	Ile	Glu	Asn	Cys	Glu	Tyr	Ala	Phe	Asn	Leu	Lys	Lys	Asp	Glu	Val	
145				150						155				160		
Cys	Val	Asn	Pro	Tyr	His	Tyr	Gln	Arg	Val	Glu	Thr	Pro	Val	Leu	Pro	

				165					170					175
Pro	Val	Leu	Val	Pro	Arg	His	Thr	Glu	Ile	Leu	Thr	Glu	Leu	Pro
			180						185				190	
Leu	Asp	Asp	Tyr	Thr	His	Ser	Ile	Pro	Glu	Asn	Thr	Asn	Phe	Pro
		195					200					205		
Gly	Ile	Glu	Pro	Gln	Ser	Asn	Tyr	Ile	Pro	Glu	Thr	Pro	Pro	Pro
	210					215					220			
Tyr	Ile	Ser	Glu	Asp	Gly	Glu	Thr	Ser	Asp	Gln	Gln	Leu	Asn	Gln
225					230					235				240
Met	Asp	Thr	Gly	Ser	Pro	Ala	Glu	Leu	Ser	Pro	Thr	Thr	Leu	Ser
				245					250				255	
Val	Asn	His	Ser	Leu	Asp	Leu	Gln	Pro	Val	Thr	Tyr	Ser	Glu	Pro
		260					265						270	
Phe	Trp	Cys	Ser	Ile	Ala	Tyr	Tyr	Glu	Leu	Asn	Gln	Arg	Val	Gly
	275						280					285		
Thr	Phe	His	Ala	Ser	Gln	Pro	Ser	Leu	Thr	Val	Asp	Gly	Phe	Thr
	290				295						300			
Pro	Ser	Asn	Ser	Glu	Arg	Phe	Cys	Leu	Gly	Leu	Leu	Ser	Asn	Val
305					310					315				320
Arg	Asn	Ala	Thr	Val	Glu	Met	Thr	Arg	Arg	His	Ile	Gly	Arg	Gly
				325					330					335
Arg	Leu	Tyr	Tyr	Ile	Gly	Gly	Glu	Val	Phe	Ala	Glu	Cys	Leu	Ser
		340						345					350	
Ser	Ala	Ile	Phe	Val	Gln	Ser	Pro	Asn	Cys	Asn	Gln	Arg	Tyr	Gly
	355					360						365		
His	Pro	Ala	Thr	Val	Cys	Lys	Ile	Pro	Pro	Gly	Cys	Asn	Leu	Lys
	370					375					380			
Phe	Asn	Asn	Gln	Glu	Phe	Ala	Ala	Leu	Leu	Ala	Gln	Ser	Val	Asn
385				390					395					400
Gly	Phe	Glu	Ala	Val	Tyr	Gln	Leu	Thr	Arg	Met	Cys	Thr	Ile	Arg
		405						410					415	
Ser	Phe	Val	Lys	Gly	Trp	Gly	Ala	Glu	Tyr	Arg	Arg	Gln	Thr	Val
		420						425					430	
Ser	Thr	Pro	Cys	Trp	Ile	Glu	Leu	His	Leu	Asn	Gly	Pro	Leu	Gln
	435					440						445		
Leu	Asp	Lys	Val	Leu	Thr	Gln	Met	Gly	Ser	Pro	Ser	Val	Arg	Cys
	450					455					460			
Ser	Met	Ser	Trp	Val	Pro	Arg	Ala	Arg	Asp	Pro	Pro	Val	Ala	Thr
465					470				475					480
Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu
		485						490					495	
Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly
		500						505					510	
Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile
	515						520					525		
Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr
	530					535					540			
Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys
545				550					555					560
His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu
				565					570					575
Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu
		580						585					590	
Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly
	595					600					605			
Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr
	610					615					620			
Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn

```

625          630          635          640
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
          645          650          655
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
          660          665          670
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
          675          680          685
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
          690          695          700
Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
705          710          715

```

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2397 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2394
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

```

ATG GAC AAT ATG TCT ATT ACG AAT ACA CCA ACA AGT AAT GAT GCC TGT      48
Met Asp Asn Met Ser Ile Thr Asn Thr Pro Thr Ser Asn Asp Ala Cys
  1              5              10              15

CTG AGC ATT GTG CAT AGT TTG ATG TGC CAT AGA CAA GGT GGA GAG AGT      96
Leu Ser Ile Val His Ser Leu Met Cys His Arg Gln Gly Gly Glu Ser
      20              25              30

GAA ACA TTT GCA AAA AGA GCA ATT GAA AGT TTG GTA AAG AAG CTG AAG      144
Glu Thr Phe Ala Lys Arg Ala Ile Glu Ser Leu Val Lys Lys Leu Lys
      35              40              45

GAG AAA AAA GAT GAA TTG GAT TCT TTA ATA ACA GCT ATA ACT ACA AAT      192
Glu Lys Lys Asp Glu Leu Asp Ser Leu Ile Thr Ala Ile Thr Thr Asn
      50              55              60

GGA GCT CAT CCT AGT AAA TGT GTT ACC ATA CAG AGA ACA TTG GAT GGG      240
Gly Ala His Pro Ser Lys Cys Val Thr Ile Gln Arg Thr Leu Asp Gly
      65              70              75              80

AGG CTT CAG GTG GCT GGT CGG AAA GGA TTT CCT CAT GTG ATC TAT GCC      288
Arg Leu Gln Val Ala Gly Arg Lys Gly Phe Pro His Val Ile Tyr Ala
      85              90              95

CGT CTC TGG AGG TGG CCT GAT CTT CAC AAA AAT GAA CTA AAA CAT GTT      336
Arg Leu Trp Arg Trp Pro Asp Leu His Lys Asn Glu Leu Lys His Val
      100              105              110

AAA TAT TGT CAG TAT GCG TTT GAC TTA AAA TGT GAT AGT GTC TGT GTG      384

```

Lys Tyr Cys Gln Tyr Ala Phe Asp Leu Lys Cys Asp Ser Val Cys Val	
115 120 125	
AAT CCA TAT CAC TAC GAA CGA GTT GTA TCA CCT GGA ATT GAT CTC TCA	432
Asn Pro Tyr His Tyr Glu Arg Val Val Ser Pro Gly Ile Asp Leu Ser	
130 135 140	
GGA TTA ACA CTG CAG AGT AAT GCT CCA TCA AGT ATG ATG GTG AAG GAT	480
Gly Leu Thr Leu Gln Ser Asn Ala Pro Ser Ser Met Met Val Lys Asp	
145 150 155 160	
GAA TAT GTG CAT GAC TTT GAG GGA CAG CCA TCG TTG TCC ACT GAA GGA	528
Glu Tyr Val His Asp Phe Glu Gly Gln Pro Ser Leu Ser Thr Glu Gly	
165 170 175	
CAT TCA ATT CAA ACC ATC CAG CAT CCA CCA AGT AAT CGT GCA TCG ACA	576
His Ser Ile Gln Thr Ile Gln His Pro Pro Ser Asn Arg Ala Ser Thr	
180 185 190	
GAG ACA TAC AGC ACC CCA GCT CTG TTA GCC CCA TCT GAG TCT AAT GCT	624
Glu Thr Tyr Ser Thr Pro Ala Leu Leu Ala Pro Ser Glu Ser Asn Ala	
195 200 205	
ACC AGC ACT GCC AAC TTT CCC AAC ATT CCT GTG GCT TCC ACA AGT CAG	672
Thr Ser Thr Ala Asn Phe Pro Asn Ile Pro Val Ala Ser Thr Ser Gln	
210 215 220	
CCT GCC AGT ATA CTG GGG GGC AGC CAT AGT GAA GGA CTG TTG CAG ATA	720
Pro Ala Ser Ile Leu Gly Gly Ser His Ser Glu Gly Leu Leu Gln Ile	
225 230 235 240	
GCA TCA GGG CCT CAG CCA GGA CAG CAG CAG AAT GGA TTT ACT GGT CAG	768
Ala Ser Gly Pro Gln Pro Gly Gln Gln Gln Asn Gly Phe Thr Gly Gln	
245 250 255	
CCA GCT ACT TAC CAT CAT AAC AGC ACT ACC ACC TGG ACT GGA AGT AGG	816
Pro Ala Thr Tyr His His Asn Ser Thr Thr Thr Trp Thr Gly Ser Arg	
260 265 270	
ACT GCA CCA TAC ACA CCT AAT TTG CCT CAC CAC CAA AAC GGC CAT CTT	864
Thr Ala Pro Tyr Thr Pro Asn Leu Pro His His Gln Asn Gly His Leu	
275 280 285	
CAG CAC CAC CCG CCT ATG CCG CCC CAT CCC GGA CAT TAC TGG CCT GTT	912
Gln His His Pro Pro Met Pro Pro His Pro Gly His Tyr Trp Pro Val	
290 295 300	
CAC AAT GAG CTT GCA TTC CAG CCT CCC ATT TCC AAT CAT CCT GCT CCT	960
His Asn Glu Leu Ala Phe Gln Pro Pro Ile Ser Asn His Pro Ala Pro	
305 310 315 320	
GAG TAT TGG TGT TCC ATT GCT TAC TTT GAA ATG GAT GTT CAG GTA GGA	1008
Glu Tyr Trp Cys Ser Ile Ala Tyr Phe Glu Met Asp Val Gln Val Gly	
325 330 335	
GAG ACA TTT AAG GTT CCT TCA AGC TGC CCT ATT GTT ACT GTT GAT GGA	1056
Glu Thr Phe Lys Val Pro Ser Ser Cys Pro Ile Val Thr Val Asp Gly	
340 345 350	

TAC GTG GAC CCT TCT GGA GGA GAT CGC TTT TGT TTG GGT CAA CTC TCC Tyr Val Asp Pro Ser Gly Gly Asp Arg Phe Cys Leu Gly Gln Leu Ser 355 360 365	1104
AAT GTC CAC AGG ACA GAA GCC ATT GAG AGA GCA AGG TTG CAC ATA GGC Asn Val His Arg Thr Glu Ala Ile Glu Arg Ala Arg Leu His Ile Gly 370 375 380	1152
AAA GGT GTG CAG TTG GAA TGT AAA GGT GAA GGT GAT GTT TGG GTC AGG Lys Gly Val Gln Leu Glu Cys Lys Gly Glu Gly Asp Val Trp Val Arg 385 390 395 400	1200
TGC CTT AGT GAC CAC GCG GTC TTT GTA CAG AGT TAC TAC TTA GAC AGA Cys Leu Ser Asp His Ala Val Phe Val Gln Ser Tyr Tyr Leu Asp Arg 405 410 415	1248
GAA GCT GGG CGT GCA CCT GGA GAT GCT GTT CAT AAG ATC TAC CCA AGT Glu Ala Gly Arg Ala Pro Gly Asp Ala Val His Lys Ile Tyr Pro Ser 420 425 430	1296
GCA TAT ATA AAG GTC TTT GAT TTG CGT CAG TGT CAT CGA CAG ATG CAG Ala Tyr Ile Lys Val Phe Asp Leu Arg Gln Cys His Arg Gln Met Gln 435 440 445	1344
CAG CAG GCG GCT ACT GCA CAA GCT GCA GCA GCT GCC CAG GCA GCA GCC Gln Gln Ala Ala Thr Ala Gln Ala Ala Ala Ala Gln Ala Ala Ala 450 455 460	1392
GTG GCA GGA AAC ATC CCT GGC CCA GGA TCA GTA GGT GGA ATA GCT CCA Val Ala Gly Asn Ile Pro Gly Pro Gly Ser Val Gly Gly Ile Ala Pro 465 470 475 480	1440
GCT ATC AGT CTG TCA GCT GCT GCT GGA ATT GGT GTT GAT GAC CTT CGT Ala Ile Ser Leu Ser Ala Ala Ala Gly Ile Gly Val Asp Asp Leu Arg 485 490 495	1488
CGC TTA TGC ATA CTC AGG ATG AGT TTT GTG AAA GGC TGG GGA CCG GAT Arg Leu Cys Ile Leu Arg Met Ser Phe Val Lys Gly Trp Gly Pro Asp 500 505 510	1536
TAC CCA AGA CAG AGC ATC AAA GAA ACA CCT TGC TGG ATT GAA ATT CAC Tyr Pro Arg Gln Ser Ile Lys Glu Thr Pro Cys Trp Ile Glu Ile His 515 520 525	1584
TTA CAC CGG GCC CTC CAG CTC CTA GAC GAA GTA CTT CAT ACC ATG CCG Leu His Arg Ala Leu Gln Leu Leu Asp Glu Val Leu His Thr Met Pro 530 535 540	1632
ATT GCA GAC CCA CAA CCT TTA GAC TGG GAT CCA CCG GTC GCC ACC ATG Ile Ala Asp Pro Gln Pro Leu Asp Trp Asp Pro Pro Val Ala Thr Met 545 550 555 560	1680
GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val 565 570 575	1728
GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG	1776

Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu			
580	585	590	
GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC	1824		
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys			
595	600	605	
ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG	1872		
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu			
610	615	620	
ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG	1920		
Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln			
625	630	635	640
CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC	1968		
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg			
645	650	655	
ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG	2016		
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val			
660	665	670	
AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC	2064		
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile			
675	680	685	
GAC TTC AAG GAG GAC GGC AAC ATC CTG GGC CAC AAG CTG GAG TAC AAC	2112		
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn			
690	695	700	
TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC	2160		
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly			
705	710	715	720
ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG	2208		
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val			
725	730	735	
CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC	2256		
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro			
740	745	750	
GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC	2304		
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser			
755	760	765	
AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG	2352		
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val			
770	775	780	
ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA	2397		
Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys			
785	790	795	

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 798 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

```

Met Asp Asn Met Ser Ile Thr Asn Thr Pro Thr Ser Asn Asp Ala Cys
 1          5          10          15
Leu Ser Ile Val His Ser Leu Met Cys His Arg Gln Gly Gly Glu Ser
          20          25          30
Glu Thr Phe Ala Lys Arg Ala Ile Glu Ser Leu Val Lys Lys Leu Lys
          35          40          45
Glu Lys Lys Asp Glu Leu Asp Ser Leu Ile Thr Ala Ile Thr Thr Asn
          50          55          60
Gly Ala His Pro Ser Lys Cys Val Thr Ile Gln Arg Thr Leu Asp Gly
65          70          75          80
Arg Leu Gln Val Ala Gly Arg Lys Gly Phe Pro His Val Ile Tyr Ala
          85          90          95
Arg Leu Trp Arg Trp Pro Asp Leu His Lys Asn Glu Leu Lys His Val
          100          105          110
Lys Tyr Cys Gln Tyr Ala Phe Asp Leu Lys Cys Asp Ser Val Cys Val
          115          120          125
Asn Pro Tyr His Tyr Glu Arg Val Val Ser Pro Gly Ile Asp Leu Ser
          130          135          140
Gly Leu Thr Leu Gln Ser Asn Ala Pro Ser Ser Met Met Val Lys Asp
          145          150          155          160
Glu Tyr Val His Asp Phe Glu Gly Gln Pro Ser Leu Ser Thr Glu Gly
          165          170          175
His Ser Ile Gln Thr Ile Gln His Pro Pro Ser Asn Arg Ala Ser Thr
          180          185          190
Glu Thr Tyr Ser Thr Pro Ala Leu Leu Ala Pro Ser Glu Ser Asn Ala
          195          200          205
Thr Ser Thr Ala Asn Phe Pro Asn Ile Pro Val Ala Ser Thr Ser Gln
          210          215          220
Pro Ala Ser Ile Leu Gly Gly Ser His Ser Glu Gly Leu Leu Gln Ile
          225          230          235          240
Ala Ser Gly Pro Gln Pro Gly Gln Gln Gln Asn Gly Phe Thr Gly Gln
          245          250          255
Pro Ala Thr Tyr His His Asn Ser Thr Thr Thr Trp Thr Gly Ser Arg
          260          265          270
Thr Ala Pro Tyr Thr Pro Asn Leu Pro His His Gln Asn Gly His Leu
          275          280          285
Gln His His Pro Pro Met Pro Pro His Pro Gly His Tyr Trp Pro Val
          290          295          300
His Asn Glu Leu Ala Phe Gln Pro Pro Ile Ser Asn His Pro Ala Pro
          305          310          315          320
Glu Tyr Trp Cys Ser Ile Ala Tyr Phe Glu Met Asp Val Gln Val Gly
          325          330          335
Glu Thr Phe Lys Val Pro Ser Ser Cys Pro Ile Val Thr Val Asp Gly
          340          345          350
Tyr Val Asp Pro Ser Gly Gly Asp Arg Phe Cys Leu Gly Gln Leu Ser
          355          360          365
Asn Val His Arg Thr Glu Ala Ile Glu Arg Ala Arg Leu His Ile Gly

```

370		375		380
Lys Gly Val Gln Leu Glu Cys Lys Gly Glu Gly Asp Val Trp Val Arg				
385		390		395
Cys Leu Ser Asp His Ala Val Phe Val Gln Ser Tyr Tyr Leu Asp Arg				400
	405		410	415
Glu Ala Gly Arg Ala Pro Gly Asp Ala Val His Lys Ile Tyr Pro Ser				
	420		425	430
Ala Tyr Ile Lys Val Phe Asp Leu Arg Gln Cys His Arg Gln Met Gln				
	435	440		445
Gln Gln Ala Ala Thr Ala Gln Ala Ala Ala Ala Gln Ala Ala Ala				
	450	455		460
Val Ala Gly Asn Ile Pro Gly Pro Gly Ser Val Gly Gly Ile Ala Pro				
465		470		475
Ala Ile Ser Leu Ser Ala Ala Ala Gly Ile Gly Val Asp Asp Leu Arg				480
	485		490	495
Arg Leu Cys Ile Leu Arg Met Ser Phe Val Lys Gly Trp Gly Pro Asp				
	500		505	510
Tyr Pro Arg Gln Ser Ile Lys Glu Thr Pro Cys Trp Ile Glu Ile His				
	515	520		525
Leu His Arg Ala Leu Gln Leu Leu Asp Glu Val Leu His Thr Met Pro				
	530	535		540
Ile Ala Asp Pro Gln Pro Leu Asp Trp Asp Pro Pro Val Ala Thr Met				
545		550		555
Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val				
	565		570	575
Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu				
	580		585	590
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys				
	595	600		605
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu				
	610	615		620
Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln				
625		630		635
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg				
	645		650	655
Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val				
	660		665	670
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile				
	675	680		685
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn				
	690	695		700
Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly				
705		710		715
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val				
	725		730	735
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro				
	740		745	750
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser				
	755	760		765
Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val				
	770	775		780
Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys				
785		790		795

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3138 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 1...3135

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

ATG GCG GGC TGG ATC CAG GCC CAG CAG CTG CAG GGA GAC GCG CTG CGC	48
Met Ala Gly Trp Ile Gln Ala Gln Gln Leu Gln Gly Asp Ala Leu Arg	
1 5 10 15	
CAG ATG CAG GTG CTG TAC GGC CAG CAC TTC CCC ATC GAG GTC CGG CAC	96
Gln Met Gln Val Leu Tyr Gly Gln His Phe Pro Ile Glu Val Arg His	
20 25 30	
TAC TTG GCC CAG TGG ATT GAG AGC CAG CCA TGG GAT GCC ATT GAC TTG	144
Tyr Leu Ala Gln Trp Ile Glu Ser Gln Pro Trp Asp Ala Ile Asp Leu	
35 40 45	
GAC AAT CCC CAG GAC AGA GCC CAA GCC ACC CAG CTC CTG GAG GGC CTG	192
Asp Asn Pro Gln Asp Arg Ala Gln Ala Thr Gln Leu Leu Glu Gly Leu	
50 55 60	
GTG CAG GAG CTG CAG AAG AAG GCG GAG CAC CAG GTG GGG GAA GAT GGG	240
Val Gln Glu Leu Gln Lys Lys Ala Glu His Gln Val Gly Glu Asp Gly	
65 70 75 80	
TTT TTA CTG AAG ATC AAG CTG GGG CAC TAC GCC ACG CAG CTC CAG AAA	288
Phe Leu Leu Lys Ile Lys Leu Gly His Tyr Ala Thr Gln Leu Gln Lys	
85 90 95	
ACA TAT GAC CGC TGC CCC CTG GAG CTG GTC CGC TGC ATC CGG CAC ATT	336
Thr Tyr Asp Arg Cys Pro Leu Glu Leu Val Arg Cys Ile Arg His Ile	
100 105 110	
CTG TAC AAT GAA CAG AGG CTG GTC CGA GAA GCC AAC AAT TGC AGC TCT	384
Leu Tyr Asn Glu Gln Arg Leu Val Arg Glu Ala Asn Asn Cys Ser Ser	
115 120 125	
CCG GCT GGG ATC CTG GTT GAC GCC ATG TCC CAG AAG CAC CTT CAG ATC	432
Pro Ala Gly Ile Leu Val Asp Ala Met Ser Gln Lys His Leu Gln Ile	
130 135 140	
AAC CAG ACA TTT GAG GAG CTG CGA CTG GTC ACG CAG GAC ACA GAG AAT	480
Asn Gln Thr Phe Glu Glu Leu Arg Leu Val Thr Gln Asp Thr Glu Asn	
145 150 155 160	
GAG CTG AAG AAA CTG CAG CAG ACT CAG GAG TAC TTC ATC ATC CAG TAC	528
Glu Leu Lys Lys Leu Gln Gln Thr Gln Glu Tyr Phe Ile Ile Gln Tyr	
165 170 175	
CAG GAG AGC CTG AGC ATC CAA GCT CAG TTT GCC CAG CTG GCC CAG CTG	576

Gln Glu Ser Leu Arg Ile Gln Ala Gln Phe Ala Gln Leu Ala Gln Leu	
180 185 190	
AGC CCC CAG GAG CGT CTG AGC CGG GAG ACG GCC CTC CAG CAG AAG CAG	624
Ser Pro Gln Glu Arg Leu Ser Arg Glu Thr Ala Leu Gln Gln Lys Gln	
195 200 205	
GTG TCT CTG GAG GCC TGG TTG CAG CGT GAG GCA CAG ACA CTG CAG CAG	672
Val Ser Leu Glu Ala Trp Leu Gln Arg Glu Ala Gln Thr Leu Gln Gln	
210 215 220	
TAC CGC GTG GAG CTG GCC GAG AAG CAC CAG AAG ACC CTG CAG CTG CTG	720
Tyr Arg Val Glu Leu Ala Glu Lys His Gln Lys Thr Leu Gln Leu Leu	
225 230 235 240	
CGG AAG CAG CAG ACC ATC ATC CTG GAT GAC GAG CTG ATC CAG TGG AAG	768
Arg Lys Gln Gln Thr Ile Ile Leu Asp Asp Glu Leu Ile Gln Trp Lys	
245 250 255	
CGG CGG CAG CAG CTG GCC GGG AAC GGC GGG CCC CCC GAG GGC AGC CTG	816
Arg Arg Gln Gln Leu Ala Gly Asn Gly Gly Pro Pro Glu Gly Ser Leu	
260 265 270	
GAC GTG CTA CAG TCC TGG TGT GAG AAG TTG GCC GAG ATC ATC TGG CAG	864
Asp Val Leu Gln Ser Trp Cys Glu Lys Leu Ala Glu Ile Ile Trp Gln	
275 280 285	
AAC CGG CAG CAG ATC CGC AGG GCT GAG CAC CTC TGC CAG CAG CTG CCC	912
Asn Arg Gln Gln Ile Arg Arg Ala Glu His Leu Cys Gln Gln Leu Pro	
290 295 300	
ATC CCC GGC CCA GTG GAG GAG ATG CTG GCC GAG GTC AAC GCC ACC ATC	960
Ile Pro Gly Pro Val Glu Glu Met Leu Ala Glu Val Asn Ala Thr Ile	
305 310 315 320	
ACG GAC ATT ATC TCA GCC CTG GTG ACC AGC ACA TTC ATC ATT GAG AAG	1008
Thr Asp Ile Ile Ser Ala Leu Val Thr Ser Thr Phe Ile Ile Glu Lys	
325 330 335	
CAG CCT CCT CAG GTC CTG AAG ACC CAG ACC AAG TTT GCA GCC ACC GTA	1056
Gln Pro Pro Gln Val Leu Lys Thr Gln Thr Lys Phe Ala Ala Thr Val	
340 345 350	
CGC CTG CTG GTG GGC GGG AAG CTG AAC GTG CAC ATG AAT CCC CCC CAG	1104
Arg Leu Leu Val Gly Gly Lys Leu Asn Val His Met Asn Pro Pro Gln	
355 360 365	
GTG AAG GCC ACC ATC ATC AGT GAG CAG CAG GCC AAG TCT CTG CTT AAA	1152
Val Lys Ala Thr Ile Ile Ser Glu Gln Gln Ala Lys Ser Leu Leu Lys	
370 375 380	
AAT GAG AAC ACC CGC AAC GAG TGC AGT GGT GAG ATC CTG AAC AAC TGC	1200
Asn Glu Asn Thr Arg Asn Glu Cys Ser Gly Glu Ile Leu Asn Asn Cys	
385 390 395 400	
TGC GTG ATG GAG TAC CAC CAA GCC ACG GGC ACC CTC AGT GCC CAC TTC	1248
Cys Val Met Glu Tyr His Gln Ala Thr Gly Thr Leu Ser Ala His Phe	
405 410 415	

AGG AAC ATG TCA CTG AAG AGG ATC AAG CGT GCT GAC CGG CGG GGT GCA	1296
Arg Asn Met Ser Leu Lys Arg Ile Lys Arg Ala Asp Arg Arg Gly Ala	
420 425 430	
GAG TCC GTG ACA GAG GAG AAG TTC ACA GTC CTG TTT GAG TCT CAG TTC	1344
Glu Ser Val Thr Glu Glu Lys Phe Thr Val Leu Phe Glu Ser Gln Phe	
435 440 445	
AGT GTT GGC AGC AAT GAG CTT GTG TTC CAG GTG AAG ACT CTG TCC CTA	1392
Ser Val Gly Ser Asn Glu Leu Val Phe Gln Val Lys Thr Leu Ser Leu	
450 455 460	
CCT GTG GTT GTC ATC GTC CAC GGC AGC CAG GAC CAC AAT GCC ACG GCT	1440
Pro Val Val Val Ile Val His Gly Ser Gln Asp His Asn Ala Thr Ala	
465 470 475 480	
ACT GTG CTG TGG GAC AAT GCC TTT GCT GAG CCG GGC AGG GTG CCA TTT	1488
Thr Val Leu Trp Asp Asn Ala Phe Ala Glu Pro Gly Arg Val Pro Phe	
485 490 495	
GCC GTG CCT GAC AAA GTG CTG TGG CCG CAG CTG TGT GAG GCG CTC AAC	1536
Ala Val Pro Asp Lys Val Leu Trp Pro Gln Leu Cys Glu Ala Leu Asn	
500 505 510	
ATG AAA TTC AAG GCC GAA GTG CAG AGC AAC CGG GGC CTG ACC AAG GAG	1584
Met Lys Phe Lys Ala Glu Val Gln Ser Asn Arg Gly Leu Thr Lys Glu	
515 520 525	
AAC CTC GTG TTC CTG GCG CAG AAA CTG TTC AAC AAC AGC AGC AGC CAC	1632
Asn Leu Val Phe Leu Ala Gln Lys Leu Phe Asn Asn Ser Ser Ser His	
530 535 540	
CTG GAG GAC TAC AGT GGC CTG TCC GTG TCC TGG TCC CAG TTC AAC AGG	1680
Leu Glu Asp Tyr Ser Gly Leu Ser Val Ser Trp Ser Gln Phe Asn Arg	
545 550 555 560	
GAG AAC TTG CCG GGC TGG AAC TAC ACC TTC TGG CAG TGG TTT GAC GGG	1728
Glu Asn Leu Pro Gly Trp Asn Tyr Thr Phe Trp Gln Trp Phe Asp Gly	
565 570 575	
GTG ATG GAG GTG TTG AAG AAG CAC CAC AAG CCC CAC TGG AAT GAT GGG	1776
Val Met Glu Val Leu Lys Lys His His Lys Pro His Trp Asn Asp Gly	
580 585 590	
GCC ATC CTA GGT TTT GTG AAT AAG CAA CAG GCC CAC GAC CTG CTC ATC	1824
Ala Ile Leu Gly Phe Val Asn Lys Gln Gln Ala His Asp Leu Leu Ile	
595 600 605	
AAC AAG CCC GAC GGG ACC TTC TTG TTG CGC TTT AGT GAC TCA GAA ATC	1872
Asn Lys Pro Asp Gly Thr Phe Leu Leu Arg Phe Ser Asp Ser Glu Ile	
610 615 620	
GGG GGC ATC ACC ATC GCC TGG AAG TTT GAC TCC CCG GAA CGC AAC CTG	1920
Gly Gly Ile Thr Ile Ala Trp Lys Phe Asp Ser Pro Glu Arg Asn Leu	
625 630 635 640	
TGG AAC CTG AAA CCA TTC ACC ACG CGG GAT TTC TCC ATC AGG TCC CTG	1968

Trp Asn Leu Lys Pro Phe Thr Thr Arg Asp Phe Ser Ile Arg Ser Leu	
645 650 655	
GCT GAC CGG CTG GGG GAC CTG AGC TAT CTC ATC TAT GTG TTT CCT GAC	2016
Ala Asp Arg Leu Gly Asp Leu Ser Tyr Leu Ile Tyr Val Phe Pro Asp	
660 665 670	
CGC CCC AAG GAT GAG GTC TTC TCC AAG TAC TAC ACT CCT GTG CTG GCT	2064
Arg Pro Lys Asp Glu Val Phe Ser Lys Tyr Tyr Thr Pro Val Leu Ala	
675 680 685	
AAA GCT GTT GAT GGA TAT GTG AAA CCA CAG ATC AAG CAA GTG GTC CCT	2112
Lys Ala Val Asp Gly Tyr Val Lys Pro Gln Ile Lys Gln Val Val Pro	
690 695 700	
GAG TTT GTG AAT GCA TCT GCA GAT GCT GGG GGC AGC AGC GCC ACG TAC	2160
Glu Phe Val Asn Ala Ser Ala Asp Ala Gly Gly Ser Ser Ala Thr Tyr	
705 710 715 720	
ATG GAC CAG GCC CCC TCC CCA GCT GTG TGC CCC CAG GCT CCC TAT AAC	2208
Met Asp Gln Ala Pro Ser Pro Ala Val Cys Pro Gln Ala Pro Tyr Asn	
725 730 735	
ATG TAC CCA CAG AAC CCT GAC CAT GTA CTC GAT CAG GAT GGA GAA TTC	2256
Met Tyr Pro Gln Asn Pro Asp His Val Leu Asp Gln Asp Gly Glu Phe	
740 745 750	
GAC CTG GAT GAG ACC ATG GAT GTG GCC AGG CAC GTG GAG GAA CTC TTA	2304
Asp Leu Asp Glu Thr Met Asp Val Ala Arg His Val Glu Glu Leu Leu	
755 760 765	
CGC CGA CCA ATG GAC AGT CTT GAC TCC CGC CTC TCG CCC CCT GCC GGT	2352
Arg Arg Pro Met Asp Ser Leu Asp Ser Arg Leu Ser Pro Pro Ala Gly	
770 775 780	
CTT TTC ACC TCT GCC AGA GGC TCC CTC TCA TGG GTA CCG CGG GCC CGG	2400
Leu Phe Thr Ser Ala Arg Gly Ser Leu Ser Trp Val Pro Arg Ala Arg	
785 790 795 800	
GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC	2448
Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr	
805 810 815	
GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC	2496
Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His	
820 825 830	
AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG	2544
Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys	
835 840 845	
CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG	2592
Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp	
850 855 860	
CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC	2640
Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg	
865 870 875 880	

TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC	2688
Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro	
885 890 895	
GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC	2736
Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn	
900 905 910	
TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC	2784
Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn	
915 920 925	
CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG	2832
Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu	
930 935 940	
GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG	2880
Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met	
945 950 955 960	
GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC	2928
Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His	
965 970 975	
AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC	2976
Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn	
980 985 990	
ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG	3024
Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu	
995 1000 1005	
AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC	3072
Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His	
1010 1015 1020	
ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG	3120
Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met	
1025 1030 1035 1040	
GAC GAG CTG TAC AAG TAA	3138
Asp Glu Leu Tyr Lys	
1045	

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1045 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Met Ala Gly Trp Ile Gln Ala Gln Gln Leu Gln Gly Asp Ala Leu Arg
 1 5 10 15
 Gln Met Gln Val Leu Tyr Gly Gln His Phe Pro Ile Glu Val Arg His
 20 25 30
 Tyr Leu Ala Gln Trp Ile Glu Ser Gln Pro Trp Asp Ala Ile Asp Leu
 35 40 45
 Asp Asn Pro Gln Asp Arg Ala Gln Ala Thr Gln Leu Leu Glu Gly Leu
 50 55 60
 Val Gln Glu Leu Gln Lys Lys Ala Glu His Gln Val Gly Glu Asp Gly
 65 70 75 80
 Phe Leu Leu Lys Ile Lys Leu Gly His Tyr Ala Thr Gln Leu Gln Lys
 85 90 95
 Thr Tyr Asp Arg Cys Pro Leu Glu Leu Val Arg Cys Ile Arg His Ile
 100 105 110
 Leu Tyr Asn Glu Gln Arg Leu Val Arg Glu Ala Asn Asn Cys Ser Ser
 115 120 125
 Pro Ala Gly Ile Leu Val Asp Ala Met Ser Gln Lys His Leu Gln Ile
 130 135 140
 Asn Gln Thr Phe Glu Glu Leu Arg Leu Val Thr Gln Asp Thr Glu Asn
 145 150 155 160
 Glu Leu Lys Lys Leu Gln Gln Thr Gln Glu Tyr Phe Ile Ile Gln Tyr
 165 170 175
 Gln Glu Ser Leu Arg Ile Gln Ala Gln Phe Ala Gln Leu Ala Gln Leu
 180 185 190
 Ser Pro Gln Glu Arg Leu Ser Arg Glu Thr Ala Leu Gln Gln Lys Gln
 195 200 205
 Val Ser Leu Glu Ala Trp Leu Gln Arg Glu Ala Gln Thr Leu Gln Gln
 210 215 220
 Tyr Arg Val Glu Leu Ala Glu Lys His Gln Lys Thr Leu Gln Leu Leu
 225 230 235 240
 Arg Lys Gln Gln Thr Ile Ile Leu Asp Asp Glu Leu Ile Gln Trp Lys
 245 250 255
 Arg Arg Gln Gln Leu Ala Gly Asn Gly Gly Pro Pro Glu Gly Ser Leu
 260 265 270
 Asp Val Leu Gln Ser Trp Cys Glu Lys Leu Ala Glu Ile Ile Trp Gln
 275 280 285
 Asn Arg Gln Gln Ile Arg Arg Ala Glu His Leu Cys Gln Gln Leu Pro
 290 295 300
 Ile Pro Gly Pro Val Glu Glu Met Leu Ala Glu Val Asn Ala Thr Ile
 305 310 315 320
 Thr Asp Ile Ile Ser Ala Leu Val Thr Ser Thr Phe Ile Ile Glu Lys
 325 330 335
 Gln Pro Pro Gln Val Leu Lys Thr Gln Thr Lys Phe Ala Ala Thr Val
 340 345 350
 Arg Leu Leu Val Gly Gly Lys Leu Asn Val His Met Asn Pro Pro Gln
 355 360 365
 Val Lys Ala Thr Ile Ile Ser Glu Gln Gln Ala Lys Ser Leu Leu Lys
 370 375 380
 Asn Glu Asn Thr Arg Asn Glu Cys Ser Gly Glu Ile Leu Asn Asn Cys
 385 390 395 400
 Cys Val Met Glu Tyr His Gln Ala Thr Gly Thr Leu Ser Ala His Phe
 405 410 415
 Arg Asn Met Ser Leu Lys Arg Ile Lys Arg Ala Asp Arg Arg Gly Ala
 420 425 430
 Glu Ser Val Thr Glu Glu Lys Phe Thr Val Leu Phe Glu Ser Gln Phe
 435 440 445
 Ser Val Gly Ser Asn Glu Leu Val Phe Gln Val Lys Thr Leu Ser Leu

450	455	460
Pro Val Val Val Ile Val His Gly Ser Gln Asp His Asn Ala Thr Ala		
465	470	475
Thr Val Leu Trp Asp Asn Ala Phe Ala Glu Pro Gly Arg Val Pro Phe		480
	485	490
		495
Ala Val Pro Asp Lys Val Leu Trp Pro Gln Leu Cys Glu Ala Leu Asn		
	500	505
		510
Met Lys Phe Lys Ala Glu Val Gln Ser Asn Arg Gly Leu Thr Lys Glu		
	515	520
		525
Asn Leu Val Phe Leu Ala Gln Lys Leu Phe Asn Asn Ser Ser Ser His		
	530	535
		540
Leu Glu Asp Tyr Ser Gly Leu Ser Val Ser Trp Ser Gln Phe Asn Arg		
545	550	555
		560
Glu Asn Leu Pro Gly Trp Asn Tyr Thr Phe Trp Gln Trp Phe Asp Gly		
	565	570
		575
Val Met Glu Val Leu Lys Lys His His Lys Pro His Trp Asn Asp Gly		
	580	585
		590
Ala Ile Leu Gly Phe Val Asn Lys Gln Gln Ala His Asp Leu Leu Ile		
	595	600
		605
Asn Lys Pro Asp Gly Thr Phe Leu Leu Arg Phe Ser Asp Ser Glu Ile		
	610	615
		620
Gly Gly Ile Thr Ile Ala Trp Lys Phe Asp Ser Pro Glu Arg Asn Leu		
625	630	635
		640
Trp Asn Leu Lys Pro Phe Thr Thr Arg Asp Phe Ser Ile Arg Ser Leu		
	645	650
		655
Ala Asp Arg Leu Gly Asp Leu Ser Tyr Leu Ile Tyr Val Phe Pro Asp		
	660	665
		670
Arg Pro Lys Asp Glu Val Phe Ser Lys Tyr Tyr Thr Pro Val Leu Ala		
	675	680
		685
Lys Ala Val Asp Gly Tyr Val Lys Pro Gln Ile Lys Gln Val Val Pro		
	690	695
		700
Glu Phe Val Asn Ala Ser Ala Asp Ala Gly Gly Ser Ser Ala Thr Tyr		
705	710	715
		720
Met Asp Gln Ala Pro Ser Pro Ala Val Cys Pro Gln Ala Pro Tyr Asn		
	725	730
		735
Met Tyr Pro Gln Asn Pro Asp His Val Leu Asp Gln Asp Gly Glu Phe		
	740	745
		750
Asp Leu Asp Glu Thr Met Asp Val Ala Arg His Val Glu Glu Leu Leu		
	755	760
		765
Arg Arg Pro Met Asp Ser Leu Asp Ser Arg Leu Ser Pro Pro Ala Gly		
	770	775
		780
Leu Phe Thr Ser Ala Arg Gly Ser Leu Ser Trp Val Pro Arg Ala Arg		
785	790	795
		800
Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr		
	805	810
		815
Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His		
	820	825
		830
Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys		
	835	840
		845
Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp		
	850	855
		860
Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg		
865	870	875
		880
Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro		
	885	890
		895
Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn		
	900	905
		910
Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn		

915	920	925
Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu		
930	935	940
Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met		
945	950	955
Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His		
	965	970
Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn		
	980	985
Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu		
	995	1000
Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His		
	1010	1015
Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met		
	1025	1030
Asp Glu Leu Tyr Lys		
		1035
		1040
		1045

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

TGGGATCCTC AGGCCGTGCT GCTGGCCG

28

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

GTCTCGAGGG AGCATGGGCA CCTTGCG

27

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

TGGGATCCGA GAAGTCTATA TCCCATC

27

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

TGGGATCCTT AGAAGTCTAT ATCCCATC

28

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

GTCTCGAGCC ATGAACGCCC CCGAGCGG

28

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

GTGAATTCTC GTCTGATTTC TGGCAGGAGG

30

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

GTGAATTC'TT TACGTCTGAT TTCTGGCAGG

30

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

GTCTCGAGCC ATGGACGAAC TGTTCCTCCCT CATC

34

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

GTGGATCCAA GGAGCTGATC TGA CTGAGCA G

31

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

GTGGATCCTT AGGAGCTGAT CTGACTCAGC AG

32

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

CCTCCTAAGC TTATCATGGA CCATTATGAT TC

32

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

CCTCCTGGAT CCCTGCGCAG GATGATGGTC CAG

33

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

GGATGGAAGC TTCAATGGCT GCCATCCGGA AGAACTGGT GATTG

45

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

GGATGGGGAT CCTCACAAGA CAAGGCAACC AGATTTTTC TTCCC

45

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

GGGAAGCTTC CATGAGCGAG ACGGTCATC

29

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

CCCGATCCT CAGGAGAAC CCCGCTTC

28

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

GTGAATTCGA CCATGGAGCG GCCCCCGGGG

30

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

GTGGTACCCA TTCTGTTAAC CAACTCC

27

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

GTGGTACCTC ATTCTGTAA CCACTCC

28

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

GTCTCGAGAG ATGCTGTCCC GTGGGTGG

28

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

GTGAATTCGC TTCCTTTGA GGGAACC

27

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

GTGAATTCAC TTCCTCTTGA GGGAACC

27

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

GTCTCGAGCC ATGGAGAACT TCCAAAAGG

29

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

GTGGATCCCA GAGTCGAAGA TGGGGTAC

28

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

GTGGATCCTC AGAGTCGAAG ATGGGGTAC

29

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

GTGAATTCGG CGATGCCAGA CCCC GCGGCG

30

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GTGGATCCCA GGCACAGGCA GCCTCAGCCT TC

32

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

GTGGATCCTC AGGCACAGGC AGCCTCAGCC TTC

33

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2616 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2613
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1 5 10 15

48

GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
20 25 30

96

GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile 35 40 45	144
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr 50 55 60	192
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys 65 70 75 80	240
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu 85 90 95	288
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu 100 105 110	336
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly 115 120 125	384
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr 130 135 140	432
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn 145 150 155 160	480
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser 165 170 175	528
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190	576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205	624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220	672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser 225 230 235 240	720
GGA CTC AGA TCT CGA GCT CAA GCT TCG AAT TCG GCG ATG CCA GAC CCC Gly Leu Arg Ser Arg Ala Gln Ala Ser Asn Ser Ala Met Pro Asp Pro 245 250 255	768
GCG GCG CAC CTG CCC TTC TTC TAC GGC AGC ATC TCG CGT GCC GAG GCC	816

Ala Ala His Leu Pro Phe Phe Tyr Gly Ser Ile Ser Arg Ala Glu Ala	
260 265 270	
GAG GAG CAC CTG AAG CTG GCG GGC ATG GCG GAC GGG CTC TTC CTG CTG	864
Glu Glu His Leu Lys Leu Ala Gly Met Ala Asp Gly Leu Phe Leu Leu	
275 280 285	
CGC CAG TGC CTG CGC TCG CTG GGC GGC TAT GTG CTG TCG CTC GTG CAC	912
Arg Gln Cys Leu Arg Ser Leu Gly Gly Tyr Val Leu Ser Leu Val His	
290 295 300	
GAT GTG CGC TTC CAC CAC TTT CCC ATC GAG CGC CAG CTC AAC GGC ACC	960
Asp Val Arg Phe His His Phe Pro Ile Glu Arg Gln Leu Asn Gly Thr	
305 310 315 320	
TAC GCC ATT GCC GGC GGC AAA GCG CAC TGT GGA CCG GCA GAG CTC TGC	1008
Tyr Ala Ile Ala Gly Gly Lys Ala His Cys Gly Pro Ala Glu Leu Cys	
325 330 335	
GAG TTC TAC TCG CGC GAC CCC GAC GGG CTG CCC TGC AAC CTG CGC AAG	1056
Glu Phe Tyr Ser Arg Asp Pro Asp Gly Leu Pro Cys Asn Leu Arg Lys	
340 345 350	
CCG TGC AAC CGG CCG TCG GGC CTC GAG CCG CAG CCG GGG GTC TTC GAC	1104
Pro Cys Asn Arg Pro Ser Gly Leu Glu Pro Gln Pro Gly Val Phe Asp	
355 360 365	
TGC CTG CGA GAC GCC ATG GTG CGT GAC TAC GTG CGC CAG ACG TGG AAG	1152
Cys Leu Arg Asp Ala Met Val Arg Asp Tyr Val Arg Gln Thr Trp Lys	
370 375 380	
CTG GAG GGC GAG GCC CTG GAG CAG GCC ATC ATC AGC CAG GCC CCG CAG	1200
Leu Glu Gly Glu Ala Leu Glu Gln Ala Ile Ile Ser Gln Ala Pro Gln	
385 390 395 400	
GTG GAG AAG CTC ATT GCT ACG ACG GCC CAC GAG CGG ATG CCC TGG TAC	1248
Val Glu Lys Leu Ile Ala Thr Thr Ala His Glu Arg Met Pro Trp Tyr	
405 410 415	
CAC AGC AGC CTG ACG CGT GAG GAG GCC GAG CGC AAA CTT TAC TCT GGG	1296
His Ser Ser Leu Thr Arg Glu Glu Ala Glu Arg Lys Leu Tyr Ser Gly	
420 425 430	
GCG CAG ACC GAC GGC AAG TTC CTG CTG AGG CCG CGG AAG GAG CAG GGC	1344
Ala Gln Thr Asp Gly Lys Phe Leu Leu Arg Pro Arg Lys Glu Gln Gly	
435 440 445	
ACA TAC GCC CTG TCC CTC ATC TAT GGG AAG ACG GTG TAC CAC TAC CTC	1392
Thr Tyr Ala Leu Ser Leu Ile Tyr Gly Lys Thr Val Tyr His Tyr Leu	
450 455 460	
ATC AGC CAA GAC AAG GCG GGC AAG TAC TGC ATT CCC GAG GGC ACC AAG	1440
Ile Ser Gln Asp Lys Ala Gly Lys Tyr Cys Ile Pro Glu Gly Thr Lys	
465 470 475 480	
TTT GAC ACG CTC TGG CAG CTG GTG GAG TAT CTG AAG CTG AAG GCG GAC	1488
Phe Asp Thr Leu Trp Gln Leu Val Glu Tyr Leu Lys Leu Lys Ala Asp	
485 490 495	

G3G CTC ATC TAC TGC CTG AAG GAG GCC TGC CCC AAC AGC AGT GCC AGC Gly Leu Ile Tyr Cys Leu Lys Glu Ala Cys Pro Asn Ser Ser Ala Ser 500 505 510	1536
AAC GCC TCA GGG GCT GCT GCT CCC ACA CTC CCA GCC CAC CCA TCC ACG Asn Ala Ser Gly Ala Ala Ala Pro Thr Leu Pro Ala His Pro Ser Thr 515 520 525	1584
TTG ACT CAT CCT CAG AGA CGA ATC GAC ACC CTC AAC TCA GAT GGA TAC Leu Thr His Pro Gln Arg Arg Ile Asp Thr Leu Asn Ser Asp Gly Tyr 530 535 540	1632
ACC CCT GAG CCA GCA CGC ATA ACG TCC CCA GAC AAA CCG CGG CCG ATG Thr Pro Glu Pro Ala Arg Ile Thr Ser Pro Asp Lys Pro Arg Pro Met 545 550 555 560	1680
CCC ATG GAC ACG AGC GTG TAT GAG AGC CCC TAC AGC GAC CCA GAG GAG Pro Met Asp Thr Ser Val Tyr Glu Ser Pro Tyr Ser Asp Pro Glu Glu 565 570 575	1728
CTC AAG GAC AAG AAG CTC TTC CTG AAG CGC GAT AAC CTC CTC ATA GCT Leu Lys Asp Lys Lys Leu Phe Leu Lys Arg Asp Asn Leu Leu Ile Ala 580 585 590	1776
GAC ATT GAA CTT GGC TGC GGC AAC TTT GGC TCA GTG CGC CAG GGC GTG Asp Ile Glu Leu Gly Cys Gly Asn Phe Gly Ser Val Arg Gln Gly Val 595 600 605	1824
TAC CGC ATG CGC AAG AAG CAG ATC GAC GTG GCC ATC AAG GTG CTG AAG Tyr Arg Met Arg Lys Lys Gln Ile Asp Val Ala Ile Lys Val Leu Lys 610 615 620	1872
CAG GGC ACG GAG AAG GCA GAC ACG GAA GAG ATG ATG CGC GAG GCG CAG Gln Gly Thr Glu Lys Ala Asp Thr Glu Glu Met Met Arg Glu Ala Gln 625 630 635 640	1920
ATC ATG CAC CAG CTG GAC AAC CCC TAC ATC GTG CGG CTC ATT GGC GTC Ile Met His Gln Leu Asp Asn Pro Tyr Ile Val Arg Leu Ile Gly Val 645 650 655	1968
TGC CAG GCC GAG GCC CTC ATG CTG GTC ATG GAG ATG GCT GGG GGC GGG Cys Gln Ala Glu Ala Leu Met Leu Val Met Glu Met Ala Gly Gly Gly 660 665 670	2016
CCG CTG CAC AAG TTC CTG GTC GGC AAG AGG GAG GAG ATC CCT GTG AGC Pro Leu His Lys Phe Leu Val Gly Lys Arg Glu Glu Ile Pro Val Ser 675 680 685	2064
AAT GTG GCC GAG CTG CTG CAC CAG GTG TCC ATG GGG ATG AAG TAC CTG Asn Val Ala Glu Leu Leu His Gln Val Ser Met Gly Met Lys Tyr Leu 690 695 700	2112
GAG GAG AAG AAC TTT GTG CAC CGT GAC CTG GCG GCC CGC AAC GTC CTG Glu Glu Lys Asn Phe Val His Arg Asp Leu Ala Ala Arg Asn Val Leu 705 710 715 720	2160
CTG GTT AAC CGG CAC TAC GCC AAG ATC AGC GAC TTT GGC CTC TCC AAA	2208

Leu Val Asn Arg His Tyr Ala Lys Ile Ser Asp Phe Gly Leu Ser Lys
 725 730 735
 GCA CTG GGT GCC GAC GAC AGC TAC TAC ACT GCC CGC TCA GCA GGG AAG 2256
 Ala Leu Gly Ala Asp Asp Ser Tyr Tyr Thr Ala Arg Ser Ala Gly Lys
 740 745 750
 TGG CCG CTC AAG TGG TAC GCA CCC GAA TGC ATC AAC TTC CGC AAG TTC 2304
 Trp Pro Leu Lys Trp Tyr Ala Pro Glu Cys Ile Asn Phe Arg Lys Phe
 755 760 765
 TCC AGC CGC AGC GAT GTC TGG AGC TAT GGG GTC ACC ATG TGG GAG GCC 2352
 Ser Ser Arg Ser Asp Val Trp Ser Tyr Gly Val Thr Met Trp Glu Ala
 770 775 780
 TTG TCC TAC GGC CAG AAG CCC TAC AAG AAG ATG AAA GGG CCG GAG GTC 2400
 Leu Ser Tyr Gly Gln Lys Pro Tyr Lys Lys Met Lys Gly Pro Glu Val
 785 790 795 800
 ATG GCC TTC ATC GAG CAG GGC AAG CGG ATG GAG TGC CCA CCA GAG TGT 2448
 Met Ala Phe Ile Glu Gln Gly Lys Arg Met Glu Cys Pro Pro Glu Cys
 805 810 815
 CCA CCC GAA CTG TAC GCA CTC ATG AGT GAC TGC TGG ATC TAC AAG TGG 2496
 Pro Pro Glu Leu Tyr Ala Leu Met Ser Asp Cys Trp Ile Tyr Lys Trp
 820 825 830
 GAG GAT CGC CCC GAC TTC CTG ACC GTG GAG CAG CGC ATG CGA GCC TGT 2544
 Glu Asp Arg Pro Asp Phe Leu Thr Val Glu Gln Arg Met Arg Ala Cys
 835 840 845
 TAC TAC AGC CTG GCC AGC AAG GTG GAA GGG CCC CCA GGC AGC ACA CAG 2592
 Tyr Tyr Ser Leu Ala Ser Lys Val Glu Gly Pro Pro Gly Ser Thr Gln
 850 855 860
 AAG GCT GAG GCT GCC TGT GCC TGA 2616
 Lys Ala Glu Ala Ala Cys Ala
 865 870

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 871 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10 15
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 20 25 30
 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile

35	40	45
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr		
50	55	60
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys		
65	70	75
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu		80
85	90	95
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu		
100	105	110
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly		
115	120	125
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr		
130	135	140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn		
145	150	155
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser		160
165	170	175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly		
180	185	190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu		
195	200	205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe		
210	215	220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser		
225	230	235
Gly Leu Arg Ser Arg Ala Gln Ala Ser Asn Ser Ala Met Pro Asp Pro		240
245	250	255
Ala Ala His Leu Pro Phe Phe Tyr Gly Ser Ile Ser Arg Ala Glu Ala		
260	265	270
Glu Glu His Leu Lys Leu Ala Gly Met Ala Asp Gly Leu Phe Leu Leu		
275	280	285
Arg Gln Cys Leu Arg Ser Leu Gly Gly Tyr Val Leu Ser Leu Val His		
290	295	300
Asp Val Arg Phe His His Phe Pro Ile Glu Arg Gln Leu Asn Gly Thr		
305	310	315
Tyr Ala Ile Ala Gly Gly Lys Ala His Cys Gly Pro Ala Glu Leu Cys		
325	330	335
Glu Phe Tyr Ser Arg Asp Pro Asp Gly Leu Pro Cys Asn Leu Arg Lys		
340	345	350
Pro Cys Asn Arg Pro Ser Gly Leu Glu Pro Gln Pro Gly Val Phe Asp		
355	360	365
Cys Leu Arg Asp Ala Met Val Arg Asp Tyr Val Arg Gln Thr Trp Lys		
370	375	380
Leu Glu Gly Glu Ala Leu Glu Gln Ala Ile Ile Ser Gln Ala Pro Gln		
385	390	395
Val Glu Lys Leu Ile Ala Thr Thr Ala His Glu Arg Met Pro Trp Tyr		
405	410	415
His Ser Ser Leu Thr Arg Glu Glu Ala Glu Arg Lys Leu Tyr Ser Gly		
420	425	430
Ala Gln Thr Asp Gly Lys Phe Leu Leu Arg Pro Arg Lys Glu Gln Gly		
435	440	445
Thr Tyr Ala Leu Ser Leu Ile Tyr Gly Lys Thr Val Tyr His Tyr Leu		
450	455	460
Ile Ser Gln Asp Lys Ala Gly Lys Tyr Cys Ile Pro Glu Gly Thr Lys		
465	470	475
Phe Asp Thr Leu Trp Gln Leu Val Glu Tyr Leu Lys Leu Lys Ala Asp		
485	490	495
Gly Leu Ile Tyr Cys Leu Lys Glu Ala Cys Pro Asn Ser Ser Ala Ser		

Asn	Ala	Ser	Gly	Ala	Ala	Ala	Pro	Thr	Leu	Pro	Ala	His	Pro	Ser	Thr
Leu	Thr	His	Pro	Gln	Arg	Arg	Ile	Asp	Thr	Leu	Asn	Ser	Asp	Gly	Tyr
Thr	Pro	Glu	Pro	Ala	Arg	Ile	Thr	Ser	Pro	Asp	Lys	Pro	Arg	Pro	Met
Pro	Met	Asp	Thr	Ser	Val	Tyr	Glu	Ser	Pro	Tyr	Ser	Asp	Pro	Glu	Glu
Leu	Lys	Asp	Lys	Lys	Leu	Phe	Leu	Lys	Arg	Asp	Asn	Leu	Leu	Ile	Ala
Asp	Ile	Glu	Leu	Gly	Cys	Gly	Asn	Phe	Gly	Ser	Val	Arg	Gln	Gly	Val
Tyr	Arg	Met	Arg	Lys	Lys	Gln	Ile	Asp	Val	Ala	Ile	Lys	Val	Leu	Lys
Gln	Gly	Thr	Glu	Lys	Ala	Asp	Thr	Glu	Glu	Met	Met	Arg	Glu	Ala	Gln
Ile	Met	His	Gln	Leu	Asp	Asn	Pro	Tyr	Ile	Val	Arg	Leu	Ile	Gly	Val
Cys	Gln	Ala	Glu	Ala	Leu	Met	Leu	Val	Met	Glu	Met	Ala	Gly	Gly	Gly
Pro	Leu	His	Lys	Phe	Leu	Val	Gly	Lys	Arg	Glu	Glu	Ile	Pro	Val	Ser
Asn	Val	Ala	Glu	Leu	Leu	His	Gln	Val	Ser	Met	Gly	Met	Lys	Tyr	Leu
Glu	Glu	Lys	Asn	Phe	Val	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Val	Leu
Leu	Val	Asn	Arg	His	Tyr	Ala	Lys	Ile	Ser	Asp	Phe	Gly	Leu	Ser	Lys
Ala	Leu	Gly	Ala	Asp	Asp	Ser	Tyr	Tyr	Thr	Ala	Arg	Ser	Ala	Gly	Lys
Trp	Pro	Leu	Lys	Trp	Tyr	Ala	Pro	Glu	Cys	Ile	Asn	Phe	Arg	Lys	Phe
Ser	Ser	Arg	Ser	Asp	Val	Trp	Ser	Tyr	Gly	Val	Thr	Met	Trp	Glu	Ala
Leu	Ser	Tyr	Gly	Gln	Lys	Pro	Tyr	Lys	Lys	Met	Lys	Gly	Pro	Glu	Val
Met	Ala	Phe	Ile	Glu	Gln	Gly	Lys	Arg	Met	Glu	Cys	Pro	Pro	Glu	Cys
Pro	Pro	Glu	Leu	Tyr	Ala	Leu	Met	Ser	Asp	Cys	Trp	Ile	Tyr	Lys	Trp
Glu	Asp	Arg	Pro	Asp	Phe	Leu	Thr	Val	Glu	Gln	Arg	Met	Arg	Ala	Cys
Tyr	Tyr	Ser	Leu	Ala	Ser	Lys	Val	Glu	Gly	Pro	Pro	Gly	Ser	Thr	Gln
Lys	Ala	Glu	Ala	Ala	Cys	Ala									

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2598 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 1...2595
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

ATG CCA GAC CCC GCG GCG CAC CTG CCC TTC TTC TAC GGC AGC ATC TCG	48
Met Pro Asp Pro Ala Ala His Leu Pro Phe Phe Tyr Gly Ser Ile Ser	
1 5 10 15	
CGT GCC GAG GCC GAG GAG CAC CTG AAG CTG GCG GGC ATG GCG GAC GGG	96
Arg Ala Glu Ala Glu Glu His Leu Lys Leu Ala Gly Met Ala Asp Gly	
20 25 30	
CTC TTC CTG CTG CGC CAG TGC CTG CGC TCG CTG GGC GGC TAT GTG CTG	144
Leu Phe Leu Leu Arg Gln Cys Leu Arg Ser Leu Gly Gly Tyr Val Leu	
35 40 45	
TCG CTC GTG CAC GAT GTG CGC TTC CAC CAC TTT CCC ATC GAG CGC CAG	192
Ser Leu Val His Asp Val Arg Phe His His Phe Pro Ile Glu Arg Gln	
50 55 60	
CTC AAC GGC ACC TAC GCC ATT GCC GGC GGC AAA GCG CAC TGT GGA CCG	240
Leu Asn Gly Thr Tyr Ala Ile Ala Gly Gly Lys Ala His Cys Gly Pro	
65 70 75 80	
GCA GAG CTC TGC GAG TTC TAC TCG CGC GAC CCC GAC GGC CTG CCC TGC	288
Ala Glu Leu Cys Glu Phe Tyr Ser Arg Asp Pro Asp Gly Leu Pro Cys	
85 90 95	
AAC CTG CGC AAG CCG TGC AAC CGG CCG TCG GGC CTC GAG CCG CAG CCG	336
Asn Leu Arg Lys Pro Cys Asn Arg Pro Ser Gly Leu Glu Pro Gln Pro	
100 105 110	
GGG GTC TTC GAC TGC CTG CGA GAC GCC ATG GTG CGT GAC TAC GTG CGC	384
Gly Val Phe Asp Cys Leu Arg Asp Ala Met Val Arg Asp Tyr Val Arg	
115 120 125	
CAG ACG TGG AAG CTG GAG GGC GAG GCC CTG GAG CAG GCC ATC ATC AGC	432
Gln Thr Trp Lys Leu Glu Gly Glu Ala Leu Glu Gln Ala Ile Ile Ser	
130 135 140	
CAG GCC CCG CAG GTG GAG AAG CTC ATT GCT ACG ACG GCC CAC GAG CGG	480
Gln Ala Pro Gln Val Glu Lys Leu Ile Ala Thr Thr Ala His Glu Arg	
145 150 155 160	
ATG CCC TGG TAC CAC AGC AGC CTG ACG CGT GAG GAG GCC GAG CGC AAA	528
Met Pro Trp Tyr His Ser Leu Thr Arg Glu Glu Ala Glu Arg Lys	
165 170 175	
CTT TAC TCT GGG GCG CAG ACC GAC GGC AAG TTC CTG CTG AGG CCG CGG	576
Leu Tyr Ser Gly Ala Gln Thr Asp Gly Lys Phe Leu Leu Arg Pro Arg	
180 185 190	
AAG GAG CAG GGC ACA TAC GCC CTG TCC CTC ATC TAT GGG AAG ACG GTG	624
Lys Glu Gln Gly Thr Tyr Ala Leu Ser Leu Ile Tyr Gly Lys Thr Val	
195 200 205	

TAC CAC TAC CTC ATC AGC CAA GAC AAG GCG GGC AAG TAC TGC ATT CCC Tyr His Tyr Leu Ile Ser Gln Asp Lys Ala Gly Lys Tyr Cys Ile Pro 210 215 220	672
GAG GGC ACC AAG TTT GAC ACG CTC TGG CAG CTG GTG GAG TAT CTG AAG Glu Gly Thr Lys Phe Asp Thr Leu Trp Gln Leu Val Glu Tyr Leu Lys 225 230 235 240	720
CTG AAG GCG GAC GGG CTC ATC TAC TGC CTG AAG GAG GCC TGC CCC AAC Leu Lys Ala Asp Gly Leu Ile Tyr Cys Leu Lys Glu Ala Cys Pro Asn 245 250 255	768
AGC AGT GCC AGC AAC GCC TCA GGG GCT GCT GCT CCC ACA CTC CCA GCC Ser Ser Ala Ser Asn Ala Ser Gly Ala Ala Pro Thr Leu Pro Ala 260 265 270	816
CAC CCA TCC ACG TTG ACT CAT CCT CAG AGA CGA ATC GAC ACC CTC AAC His Pro Ser Thr Leu Thr His Pro Gln Arg Arg Ile Asp Thr Leu Asn 275 280 285	864
TCA GAT GGA TAC ACC CCT GAG CCA GCA CGC ATA ACG TCC CCA GAC AAA Ser Asp Gly Tyr Thr Pro Glu Pro Ala Arg Ile Thr Ser Pro Asp Lys 290 295 300	912
CCG CGG CCG ATG CCC ATG GAC ACG AGC GTG TAT GAG AGC CCC TAC AGC Pro Arg Pro Met Pro Met Asp Thr Ser Val Tyr Glu Ser Pro Tyr Ser 305 310 315 320	960
GAC CCA GAG GAG CTC AAG GAC AAG AAG CTC TTC CTG AAG CGC GAT AAC Asp Pro Glu Glu Leu Lys Asp Lys Lys Leu Phe Leu Lys Arg Asp Asn 325 330 335	1008
CTC CTC ATA GCT GAC ATT GAA CTT GGC TGC GGC AAC TTT GGC TCA GTG Leu Leu Ile Ala Asp Ile Glu Leu Gly Cys Gly Asn Phe Gly Ser Val 340 345 350	1056
CGC CAG GGC GTG TAC CGC ATG CGC AAG AAG CAG ATC GAC GTG GCC ATC Arg Gln Gly Val Tyr Arg Met Arg Lys Lys Gln Ile Asp Val Ala Ile 355 360 365	1104
AAG GTG CTG AAG CAG GGC ACG GAG AAG GCA GAC ACG GAA GAG ATG ATG Lys Val Leu Lys Gln Gly Thr Glu Lys Ala Asp Thr Glu Glu Met Met 370 375 380	1152
CGC GAG GCG CAG ATC ATG CAC CAG CTG GAC AAC CCC TAC ATC GTG CGG Arg Glu Ala Gln Ile Met His Gln Leu Asp Asn Pro Tyr Ile Val Arg 385 390 395 400	1200
CTC ATT GGC GTC TGC CAG GCC GAG GCC CTC ATG CTG CTC ATG GAG ATG Leu Ile Gly Val Cys Gln Ala Glu Ala Leu Met Leu Val Met Glu Met 405 410 415	1248
GCT GGG GGC GGG CCG CTG CAC AAG TTC CTG GTC GGC AAG AGG GAG GAG Ala Gly Gly Gly Pro Leu His Lys Phe Leu Val Gly Lys Arg Glu Glu 420 425 430	1296
ATC CCT GTG AGC AAT GTG GCC GAG CTG CTG CAC CAG GTG TCC ATG GGG	1344

Ile	Pro	Val	Ser	Asn	Val	Ala	Glu	Leu	Leu	His	Gln	Val	Ser	Met	Gly		
	435							440					445				
ATG	AAG	TAC	CTG	GAG	GAG	AAG	AAC	TTT	GTG	CAC	CGT	GAC	CTG	GCG	GCC	1392	
Met	Lys	Tyr	Leu	Glu	Glu	Lys	Asn	Phe	Val	His	Arg	Asp	Leu	Ala	Ala		
	450					455					460						
CGC	AAC	GTC	CTG	CTG	GTT	AAC	CGG	CAC	TAC	GCC	AAG	ATC	AGC	GAC	TTT	1440	
Arg	Asn	Val	Leu	Leu	Val	Asn	Arg	His	Tyr	Ala	Lys	Ile	Ser	Asp	Phe		
	465				470					475					480		
GGC	CTC	TCC	AAA	GCA	CTG	GGT	GCC	GAC	GAC	AGC	TAC	TAC	ACT	GCC	CGC	1488	
Gly	Leu	Ser	Lys	Ala	Leu	Gly	Ala	Asp	Asp	Ser	Tyr	Tyr	Thr	Ala	Arg		
				485					490					495			
TCA	GCA	GGG	AAG	TGG	CCG	CTC	AAG	TGG	TAC	GCA	CCC	GAA	TGC	ATC	AAC	1536	
Ser	Ala	Gly	Lys	Trp	Pro	Leu	Lys	Trp	Tyr	Ala	Pro	Glu	Cys	Ile	Asn		
		500						505					510				
TTC	CGC	AAG	TTC	TCC	AGC	CGC	AGC	GAT	GTC	TGG	AGC	TAT	GGG	GTC	ACC	1584	
Phe	Arg	Lys	Phe	Ser	Ser	Arg	Ser	Asp	Val	Trp	Ser	Tyr	Gly	Val	Thr		
		515						520					525				
ATG	TGG	GAG	GCC	TTG	TCC	TAC	GGC	CAG	AAG	CCC	TAC	AAG	AAG	ATG	AAA	1632	
Met	Trp	Glu	Ala	Leu	Ser	Tyr	Gly	Gln	Lys	Pro	Tyr	Lys	Lys	Met	Lys		
	530					535					540						
GGG	CCG	GAG	GTC	ATG	GCC	TTC	ATC	GAG	CAG	GGC	AAG	CGG	ATG	GAG	TGC	1680	
Gly	Pro	Glu	Val	Met	Ala	Phe	Ile	Glu	Gln	Gly	Lys	Arg	Met	Glu	Cys		
	545				550					555				560			
CCA	CCA	GAG	TGT	CCA	CCC	GAA	CTG	TAC	GCA	CTC	ATG	AGT	GAC	TGC	TGG	1728	
Pro	Pro	Glu	Cys	Pro	Pro	Glu	Leu	Tyr	Ala	Leu	Met	Ser	Asp	Cys	Trp		
			565					570						575			
ATC	TAC	AAG	TGG	GAG	GAT	CGC	CCC	GAC	TTC	CTG	ACC	GTG	GAG	CAG	CGC	1776	
Ile	Tyr	Lys	Trp	Glu	Asp	Arg	Pro	Asp	Phe	Leu	Thr	Val	Glu	Gln	Arg		
		580						585						590			
ATG	CGA	GCC	TGT	TAC	TAC	AGC	CTG	GCC	AGC	AAG	GTG	GAA	GGG	CCC	CCA	1824	
Met	Arg	Ala	Cys	Tyr	Tyr	Ser	Leu	Ala	Ser	Lys	Val	Glu	Gly	Pro	Pro		
		595					600					605					
GGC	AGC	ACA	CAG	AAG	GCT	GAG	GCT	GCC	TGT	GCC	TGG	GAT	CCA	CCG	GTC	1872	
Gly	Ser	Thr	Gln	Lys	Ala	Glu	Ala	Ala	Cys	Ala	Trp	Asp	Pro	Pro	Val		
	610					615					620						
GCC	ACC	ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	1920	
Ala	Thr	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro		
	625				630					635				640			
ATC	CTG	GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	1968	
Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val		
			645					650						655			
TCC	GGC	GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	AAG	2016	
Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys		
		660					665						670				

TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val 675 680 685	2064
ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His 690 695 700	2112
ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val 705 710 715 720	2160
CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg 725 730 735	2208
GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu 740 745 750	2256
AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu 755 760 765	2304
GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln 770 775 780	2352
AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp 785 790 795 800	2400
GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly 805 810 815	2448
GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser 820 825 830	2496
GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu 835 840 845	2544
GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr 850 855 860	2592
AAG TAA Lys 865	2598

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 865 amino acids

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

Met	Pro	Asp	Pro	Ala	Ala	His	Leu	Pro	Phe	Phe	Tyr	Gly	Ser	Ile	Ser
1				5				10						15	
Arg	Ala	Glu	Ala	Glu	Glu	His	Leu	Lys	Leu	Ala	Gly	Met	Ala	Asp	Gly
		20						25					30		
Leu	Phe	Leu	Leu	Arg	Gln	Cys	Leu	Arg	Ser	Leu	Gly	Gly	Tyr	Val	Leu
		35					40					45			
Ser	Leu	Val	His	Asp	Val	Arg	Phe	His	His	Phe	Pro	Ile	Glu	Arg	Gln
		50				55					60				
Leu	Asn	Gly	Thr	Tyr	Ala	Ile	Ala	Gly	Gly	Lys	Ala	His	Cys	Gly	Pro
65					70					75				80	
Ala	Glu	Leu	Cys	Glu	Phe	Tyr	Ser	Arg	Asp	Pro	Asp	Gly	Leu	Pro	Cys
			85						90				95		
Asn	Leu	Arg	Lys	Pro	Cys	Asn	Arg	Pro	Ser	Gly	Leu	Glu	Pro	Gln	Pro
			100					105					110		
Gly	Val	Phe	Asp	Cys	Leu	Arg	Asp	Ala	Met	Val	Arg	Asp	Tyr	Val	Arg
		115				120					125				
Gln	Thr	Trp	Lys	Leu	Glu	Gly	Glu	Ala	Leu	Glu	Gln	Ala	Ile	Ile	Ser
		130				135					140				
Gln	Ala	Pro	Gln	Val	Glu	Lys	Leu	Ile	Ala	Thr	Thr	Ala	His	Glu	Arg
145					150					155				160	
Met	Pro	Trp	Tyr	His	Ser	Ser	Leu	Thr	Arg	Glu	Glu	Ala	Glu	Arg	Lys
			165						170					175	
Leu	Tyr	Ser	Gly	Ala	Gln	Thr	Asp	Gly	Lys	Phe	Leu	Leu	Arg	Pro	Arg
		180						185					190		
Lys	Glu	Gln	Gly	Thr	Tyr	Ala	Leu	Ser	Leu	Ile	Tyr	Gly	Lys	Thr	Val
		195				200					205				
Tyr	His	Tyr	Leu	Ile	Ser	Gln	Asp	Lys	Ala	Gly	Lys	Tyr	Cys	Ile	Pro
	210				215						220				
Glu	Gly	Thr	Lys	Phe	Asp	Thr	Leu	Trp	Gln	Leu	Val	Glu	Tyr	Leu	Lys
225				230						235				240	
Leu	Lys	Ala	Asp	Gly	Leu	Ile	Tyr	Cys	Leu	Lys	Glu	Ala	Cys	Pro	Asn
			245						250				255		
Ser	Ser	Ala	Ser	Asn	Ala	Ser	Gly	Ala	Ala	Ala	Pro	Thr	Leu	Pro	Ala
		260						265					270		
His	Pro	Ser	Thr	Leu	Thr	His	Pro	Gln	Arg	Arg	Ile	Asp	Thr	Leu	Asn
		275				280					285				
Ser	Asp	Gly	Tyr	Thr	Pro	Glu	Pro	Ala	Arg	Ile	Thr	Ser	Pro	Asp	Lys
	290					295					300				
Pro	Arg	Pro	Met	Pro	Met	Asp	Thr	Ser	Val	Tyr	Glu	Ser	Pro	Tyr	Ser
305				310						315				320	
Asp	Pro	Glu	Glu	Leu	Lys	Asp	Lys	Lys	Leu	Phe	Leu	Lys	Arg	Asp	Asn
			325						330				335		
Leu	Leu	Ile	Ala	Asp	Ile	Glu	Leu	Gly	Cys	Gly	Asn	Phe	Gly	Ser	Val
		340						345					350		
Arg	Gln	Gly	Val	Tyr	Arg	Met	Arg	Lys	Lys	Gln	Ile	Asp	Val	Ala	Ile
		355				360					365				
Lys	Val	Leu	Lys	Gln	Gly	Thr	Glu	Lys	Ala	Asp	Thr	Glu	Glu	Met	Met
	370					375					380				
Arg	Glu	Ala	Gln	Ile	Met	His	Gln	Leu	Asp	Asn	Pro	Tyr	Ile	Val	Arg

385		390		395		400									
Leu	Ile	Gly	Val	Cys	Gln	Ala	Glu	Ala	Leu	Met	Leu	Val	Met	Glu	Met
		405							410					415	
Ala	Gly	Gly	Gly	Pro	Leu	His	Lys	Phe	Leu	Val	Gly	Lys	Arg	Glu	Glu
		420						425						430	
Ile	Pro	Val	Ser	Asn	Val	Ala	Glu	Leu	Leu	His	Gln	Val	Ser	Met	Gly
		435						440						445	
Met	Lys	Tyr	Leu	Glu	Glu	Lys	Asn	Phe	Val	His	Arg	Asp	Leu	Ala	Ala
		450						455						460	
Arg	Asn	Val	Leu	Leu	Val	Asn	Arg	His	Tyr	Ala	Lys	Ile	Ser	Asp	Phe
		465						470						475	
Gly	Leu	Ser	Lys	Ala	Leu	Gly	Ala	Asp	Asp	Ser	Tyr	Tyr	Thr	Ala	Arg
				485					490						495
Ser	Ala	Gly	Lys	Trp	Pro	Leu	Lys	Trp	Tyr	Ala	Pro	Glu	Cys	Ile	Asn
			500						505					510	
Phe	Arg	Lys	Phe	Ser	Ser	Arg	Ser	Asp	Val	Trp	Ser	Tyr	Gly	Val	Thr
		515							520					525	
Met	Trp	Glu	Ala	Leu	Ser	Tyr	Gly	Gln	Lys	Pro	Tyr	Lys	Lys	Met	Lys
		530							535					540	
Gly	Pro	Glu	Val	Met	Ala	Phe	Ile	Glu	Gln	Gly	Lys	Arg	Met	Glu	Cys
		545								550				555	
Pro	Pro	Glu	Cys	Pro	Pro	Glu	Leu	Tyr	Ala	Leu	Met	Ser	Asp	Cys	Trp
				565						570					575
Ile	Tyr	Lys	Trp	Glu	Asp	Arg	Pro	Asp	Phe	Leu	Thr	Val	Glu	Gln	Arg
				580						585					590
Met	Arg	Ala	Cys	Tyr	Tyr	Ser	Leu	Ala	Ser	Lys	Val	Glu	Gly	Pro	Pro
			595					600						605	
Gly	Ser	Thr	Gln	Lys	Ala	Glu	Ala	Ala	Cys	Ala	Trp	Asp	Pro	Pro	Val
		610						615						620	
Ala	Thr	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro
					630						635				640
Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val
				645						650					655
Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys
			660						665						670
Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val
			675					680						685	
Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His
		690						695						700	
Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val
				705					710					715	
Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg
				725						730					735
Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu
				740						745					750
Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu
			755						760						765
Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln
			770						775					780	
Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp
			785						790						800
Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly
				805						810					815
Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser
				820						825					830
Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu
				835						840					845
Glu	Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr

850
Lys
865

855

860

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1635 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1632
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

ATG GAG AAC TTC CAA AAG GTG GAA AAG ATC GGA GAG GGC ACG TAC GGA	48
Met Glu Asn Phe Gln Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly	
1 5 10 15	
GTT GTG TAC AAA GCC AGA AAC AAG TTG ACG GGA GAG GTG GTG GCG CTT	96
Val Val Tyr Lys Ala Arg Asn Lys Leu Thr Gly Glu Val Val Ala Leu	
20 25 30	
AAG AAA ATC CGC CTG GAC ACT GAG ACT GAG GGT GTG CCC AGT ACT GCC	144
Lys Lys Ile Arg Leu Asp Thr Glu Thr Glu Gly Val Pro Ser Thr Ala	
35 40 45	
ATC CGA GAG ATC TCT CTG CTT AAG GAG CTT AAC CAT CCT AAT ATT GTC	192
Ile Arg Glu Ile Ser Leu Leu Lys Glu Leu Asn His Pro Asn Ile Val	
50 55 60	
AAG CTG CTG GAT GTC ATT CAC ACA GAA AAT AAA CTC TAC CTG GTT TTT	240
Lys Leu Leu Asp Val Ile His Thr Glu Asn Lys Leu Tyr Leu Val Phe	
65 70 75 80	
GAA TTT CTG CAC CAA GAT CTC AAG AAA TTC ATG GAT GCC TCT GCT CTC	288
Glu Phe Leu His Gln Asp Leu Lys Lys Phe Met Asp Ala Ser Ala Leu	
85 90 95	
ACT GGC ATT CCT CTT CCC CTC ATC AAG AGC TAT CTG TTC CAG CTG CTC	336
Thr Gly Ile Pro Leu Pro Leu Ile Lys Ser Tyr Leu Phe Gln Leu Leu	
100 105 110	
CAG GGC CTA GCT TTC TGC CAT TCT CAT CGG GTC CTC CAC CGA GAC CTT	384
Gln Gly Leu Ala Phe Cys His Ser His Arg Val Leu His Arg Asp Leu	
115 120 125	
AAA CCT CAG AAT CTG CTT ATT AAC ACA GAG GGG GCC ATC AAG CTA GCA	432
Lys Pro Gln Asn Leu Leu Ile Asn Thr Glu Gly Ala Ile Lys Leu Ala	
130 135 140	
GAC TTT GGA CTA GCC AGA GCT TTT GGA GTC CCT GTT CGT ACT TAC ACC	480

[illegible]

(2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 544 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

Met Glu Asn Phe Gln Lys Val Glu Lys Ile Gly Glu Gly Thr Tyr Gly
1 5 10 15

Val Val Tyr Lys Ala Arg Asn Lys Leu Thr Gly Glu Val Val Ala Leu
 20 25 30
 Lys Lys Ile Arg Leu Asp Thr Glu Thr Glu Gly Val Pro Ser Thr Ala
 35 40 45
 Ile Arg Glu Ile Ser Leu Leu Lys Glu Leu Asn His Pro Asn Ile Val
 50 55 60
 Lys Leu Leu Asp Val Ile His Thr Glu Asn Lys Leu Tyr Leu Val Phe
 65 70 75 80
 Glu Phe Leu His Gln Asp Leu Lys Lys Phe Met Asp Ala Ser Ala Leu
 85 90 95
 Thr Gly Ile Pro Leu Pro Leu Ile Lys Ser Tyr Leu Phe Gln Leu Leu
 100 105 110
 Gln Gly Leu Ala Phe Cys His Ser His Arg Val Leu His Arg Asp Leu
 115 120 125
 Lys Pro Gln Asn Leu Leu Ile Asn Thr Glu Gly Ala Ile Lys Leu Ala
 130 135 140
 Asp Phe Gly Leu Ala Arg Ala Phe Gly Val Pro Val Arg Thr Tyr Thr
 145 150 155 160
 His Glu Val Val Thr Leu Trp Tyr Arg Ala Pro Glu Ile Leu Leu Gly
 165 170 175
 Ser Lys Tyr Tyr Ser Thr Ala Val Asp Ile Trp Ser Leu Gly Cys Ile
 180 185 190
 Phe Ala Glu Met Val Thr Arg Arg Ala Leu Phe Pro Gly Asp Ser Glu
 195 200 205
 Ile Asp Gln Leu Phe Arg Ile Phe Arg Thr Leu Gly Thr Pro Asp Glu
 210 215 220
 Val Val Trp Pro Gly Val Thr Ser Met Pro Asp Tyr Lys Pro Ser Phe
 225 230 235 240
 Pro Lys Trp Ala Arg Gln Asp Phe Ser Lys Val Val Pro Pro Leu Asp
 245 250 255
 Glu Asp Gly Arg Ser Leu Leu Ser Gln Met Leu His Tyr Asp Pro Asn
 260 265 270
 Lys Arg Ile Ser Ala Lys Ala Ala Leu Ala His Pro Phe Phe Gln Asp
 275 280 285
 Val Thr Lys Pro Val Pro His Leu Arg Leu Trp Asp Pro Pro Val Ala
 290 295 300
 Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile
 305 310 315 320
 Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser
 325 330 335
 Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe
 340 345 350
 Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr
 355 360 365
 Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met
 370 375 380
 Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln
 385 390 395 400
 Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala
 405 410 415
 Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys
 420 425 430
 Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu
 435 440 445
 Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys
 450 455 460
 Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly
 465 470 475 480

Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp
				485					490					495	
Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala
			500					505					510		
Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu
		515					520					525			
Phe	Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys
	530					535					540				

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1635 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...1632
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	48
Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
1				5					10					15		
GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	96
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			20					25					30			
GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	AAG	TTC	ATC	144
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
			35				40					45				
TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	192
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
	50					55					60					
CTG	ACC	TAC	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	CAC	ATG	AAG	240
Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	
65				70					75						80	
CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	288
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	
				85					90					95		
CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	336
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	
			100					105					110			
GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	CTG	AAG	GGC	384
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	
			115				120					125				

ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr 130 135 140	432
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn 145 150 155 160	480
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser 165 170 175	528
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190	576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205	624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220	672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser 225 230 235 240	720
GGA CTC AGA TCT CGA GCC ATG GAG AAC TTC CAA AAG GTG GAA AAG ATC Gly Leu Arg Ser Arg Ala Met Glu Asn Phe Gln Lys Val Glu Lys Ile 245 250 255	768
GGA GAG GGC ACG TAC GGA GTT GTG TAC AAA GCC AGA AAC AAG TTG ACG Gly Glu Gly Thr Tyr Gly Val Val Tyr Lys Ala Arg Asn Lys Leu Thr 260 265 270	816
GGA GAG GTG GTG GCG CTT AAG AAA ATC CGC CTG GAC ACT GAG ACT GAG Gly Glu Val Val Ala Leu Lys Lys Ile Arg Leu Asp Thr Glu Thr Glu 275 280 285	864
GGT GTG CCC AGT ACT GCC ATC CGA GAG ATC TCT CTG CTT AAG GAG CTT Gly Val Pro Ser Thr Ala Ile Arg Glu Ile Ser Leu Leu Lys Glu Leu 290 295 300	912
AAC CAT CCT AAT ATT GTC AAG CTG CTG GAT GTC ATT CAC ACA GAA AAT Asn His Pro Asn Ile Val Lys Leu Leu Asp Val Ile His Thr Glu Asn 305 310 315 320	960
AAA CTC TAC CTG GTT TTT GAA TTT CTG CAC CAA GAT CTC AAG AAA TTC Lys Leu Tyr Leu Val Phe Glu Phe Leu His Gln Asp Leu Lys Lys Phe 325 330 335	1008
ATG GAT GCC TCT GCT CTC ACT GGC ATT CCT CTT CCC CTC ATC AAG AGC Met Asp Ala Ser Ala Leu Thr Gly Ile Pro Leu Pro Leu Ile Lys Ser 340 345 350	1056
TAT CTG TTC CAG CTG CTC CAG GGC CTA GCT TTC TGC CAT TCT CAT CGG Tyr Leu Phe Gln Leu Leu Gln Gly Leu Ala Phe Cys His Ser His Arg	1104

355	360	365	
GTC CTC CAC CGA GAC CTT AAA CCT CAG AAT CTG CTT ATT AAC ACA GAG			1152
Val Leu His Arg Asp Leu Lys Pro Gln Asn Leu Leu Ile Asn Thr Glu			
370	375	380	
GGG GCC ATC AAG CTA GCA GAC TTT GGA CTA GCC AGA GCT TTT GGA GTC			1200
Gly Ala Ile Lys Leu Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Val			
385	390	395	400
CCT GTT CGT ACT TAC ACC CAT GAG GTG GTG ACC CTG TGG TAC CGA GCT			1248
Pro Val Arg Thr Tyr Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala			
405	410	415	
CCT GAA ATC CTC CTG GGC TCG AAA TAT TAT TCC ACA GCT GTG GAC ATC			1296
Pro Glu Ile Leu Leu Gly Ser Lys Tyr Tyr Ser Thr Ala Val Asp Ile			
420	425	430	
TGG AGC CTG GGC TGC ATC TTT GCT GAG ATG GTG ACT CGC CGG GCC CTG			1344
Trp Ser Leu Gly Cys Ile Phe Ala Glu Met Val Thr Arg Arg Ala Leu			
435	440	445	
TTC CCT GGA GAT TCT GAG ATT GAC CAG CTC TTC CGG ATC TTT CGG ACT			1392
Phe Pro Gly Asp Ser Glu Ile Asp Gln Leu Phe Arg Ile Phe Arg Thr			
450	455	460	
CTG GGG ACC CCA GAT GAG GTG GTG TGG CCA GGA GTT ACT TCT ATG CCT			1440
Leu Gly Thr Pro Asp Glu Val Val Trp Pro Gly Val Thr Ser Met Pro			
465	470	475	480
GAT TAC AAG CCA AGT TTC CCC AAG TGG GCC CGG CAA GAT TTT AGT AAA			1488
Asp Tyr Lys Pro Ser Phe Pro Lys Trp Ala Arg Gln Asp Phe Ser Lys			
485	490	495	
GTT GTA CCT CCC CTG GAT GAA GAT GGA CGG AGC TTG TTA TCG CAA ATG			1536
Val Val Pro Pro Leu Asp Glu Asp Gly Arg Ser Leu Leu Ser Gln Met			
500	505	510	
CTG CAC TAC GAC CCT AAC AAG CGG ATT TCG GCC AAG GCA GCC CTG GCT			1584
Leu His Tyr Asp Pro Asn Lys Arg Ile Ser Ala Lys Ala Ala Leu Ala			
515	520	525	
CAC CCT TTC TTC CAG GAT GTG ACC AAG CCA GTA CCC CAT CTT CGA CTC T			1633
His Pro Phe Phe Gln Asp Val Thr Lys Pro Val Pro His Leu Arg Leu			
530	535	540	
GA			1635

(2) INFORMATION FOR SEQ ID NO:115:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 544 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

```

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1           5           10           15
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
          20           25           30
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
          35           40           45
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
          50           55           60
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65           70           75           80
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
          85           90           95
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
          100          105          110
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
          115          120          125
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
          130          135          140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145           150           155           160
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
          165          170          175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
          180          185          190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
          195          200          205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
          210          215          220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
225           230           235           240
Gly Leu Arg Ser Arg Ala Met Glu Asn Phe Gln Lys Val Glu Lys Ile
          245          250          255
Gly Glu Gly Thr Tyr Gly Val Val Tyr Lys Ala Arg Asn Lys Leu Thr
          260          265          270
Gly Glu Val Val Ala Leu Lys Lys Ile Arg Leu Asp Thr Glu Thr Glu
          275          280          285
Gly Val Pro Ser Thr Ala Ile Arg Glu Ile Ser Leu Leu Lys Glu Leu
          290          295          300
Asn His Pro Asn Ile Val Lys Leu Leu Asp Val Ile His Thr Glu Asn
305           310           315           320
Lys Leu Tyr Leu Val Phe Glu Phe Leu His Gln Asp Leu Lys Lys Phe
          325          330          335
Met Asp Ala Ser Ala Leu Thr Gly Ile Pro Leu Pro Leu Ile Lys Ser
          340          345          350
Tyr Leu Phe Gln Leu Leu Gln Gly Leu Ala Phe Cys His Ser His Arg
          355          360          365
Val Leu His Arg Asp Leu Lys Pro Gln Asn Leu Leu Ile Asn Thr Glu
          370          375          380
Gly Ala Ile Lys Leu Ala Asp Phe Gly Leu Ala Arg Ala Phe Gly Val
385           390           395           400
Pro Val Arg Thr Tyr Thr His Glu Val Val Thr Leu Trp Tyr Arg Ala
          405          410          415
Pro Glu Ile Leu Leu Gly Ser Lys Tyr Tyr Ser Thr Ala Val Asp Ile
          420          425          430
Trp Ser Leu Gly Cys Ile Phe Ala Glu Met Val Thr Arg Arg Ala Leu

```

435		440		445
Phe Pro Gly Asp Ser Glu Ile Asp Gln Leu Phe Arg Ile Phe Arg Thr				
450		455		460
Leu Gly Thr Pro Asp Glu Val Val Trp Pro Gly Val Thr Ser Met Pro				
465		470		475
Asp Tyr Lys Pro Ser Phe Pro Lys Trp Ala Arg Gln Asp Phe Ser Lys				
	485		490	495
Val Val Pro Pro Leu Asp Glu Asp Gly Arg Ser Leu Leu Ser Gln Met				
	500		505	510
Leu His Tyr Asp Pro Asn Lys Arg Ile Ser Ala Lys Ala Ala Leu Ala				
	515		520	525
His Pro Phe Phe Gln Asp Val Thr Lys Pro Val Pro His Leu Arg Leu				
530		535		540

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2532 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2529
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG	336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
100 105 110	

GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly 115 120 125	384
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr 130 135 140	432
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn 145 150 155 160	480
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser 165 170 175	528
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190	576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205	624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220	672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser 225 230 235 240	720
GGA CTC AGA TCT CGA GAG ATG CTG TCC CGT GGG TGG TTT CAC CGA GAC Gly Leu Arg Ser Arg Glu Met Leu Ser Arg Gly Trp Phe His Arg Asp 245 250 255	768
CTC AGT GGG CTG GAT GCA GAG ACC CTG CTC AAG GGC CGA GGT GTC CAC Leu Ser Gly Leu Asp Ala Glu Thr Leu Leu Lys Gly Arg Gly Val His 260 265 270	816
GGT AGC TTC CTG GCT CGG CCC AGT CGC AAG AAC CAG GGT GAC TTC TCG Gly Ser Phe Leu Ala Arg Pro Ser Arg Lys Asn Gln Gly Asp Phe Ser 275 280 285	864
CTC TCC GTC AGG GTG GGG GAT CAG GTG ACC CAT ATT CGG ATC CAG AAC Leu Ser Val Arg Val Gly Asp Gln Val Thr His Ile Arg Ile Gln Asn 290 295 300	912
TCA GGG GAT TTC TAT GAC CTG TAT GGA GGG GAG AAG TTT GCG ACT CTG Ser Gly Asp Phe Tyr Asp Leu Tyr Gly Gly Glu Lys Phe Ala Thr Leu 305 310 315 320	960
ACA GAG CTG GTG GAG TAC TAC ACT CAG CAG CAG GGT GTC CTG CAG GAC Thr Glu Leu Val Glu Tyr Tyr Thr Gln Gln Gln Gly Val Leu Gln Asp 325 330 335	1008
CGC GAC GGC ACC ATC ATC CAC CTC AAG TAC CCG CTG AAC TGC TCC GAT	1056

Arg Asp Gly Thr Ile Ile His Leu Lys Tyr Pro Leu Asn Cys Ser Asp	
340 345 350	
CCC ACT AGT GAG AGG TGG TAC CAT GGC CAC ATG TCT GGC GGG CAG GCA	1104
Pro Thr Ser Glu Arg Trp Tyr His Gly His Met Ser Gly Gly Gln Ala	
355 360 365	
GAG ACG CTG CTG CAG GCC AAG GGC GAG CCC TGG ACG TTT CTT GTG CGT	1152
Glu Thr Leu Leu Gln Ala Lys Gly Glu Pro Trp Thr Phe Leu Val Arg	
370 375 380	
GAG AGC CTC AGC CAG CCT GGA GAC TTC GTG CTT TCT GTG CTC AGT GAC	1200
Glu Ser Leu Ser Gln Pro Gly Asp Phe Val Leu Ser Val Leu Ser Asp	
385 390 395 400	
CAG CCC AAG GCT GGC CCA GGC TCC CCG CTC AGG GTC ACC CAC ATC AAG	1248
Gln Pro Lys Ala Gly Pro Gly Ser Pro Leu Arg Val Thr His Ile Lys	
405 410 415	
GTC ATG TGC GAG GGT GGA CGC TAC ACA GTG GGT GGT TTG GAG ACC TTC	1296
Val Met Cys Glu Gly Gly Arg Tyr Thr Val Gly Gly Leu Glu Thr Phe	
420 425 430	
GAC AGC CTC ACG GAC CTG GTA GAG CAT TTC AAG AAG ACG GGG ATT GAG	1344
Asp Ser Leu Thr Asp Leu Val Glu His Phe Lys Lys Thr Gly Ile Glu	
435 440 445	
GAG GCC TCA GGC GCC TTT GTC TAC CTG CGG CAG CCG TAC TAT GCC ACG	1392
Glu Ala Ser Gly Ala Phe Val Tyr Leu Arg Gln Pro Tyr Tyr Ala Thr	
450 455 460	
AGG GTG AAT GCG GCT GAC ATT GAG AAC CGA GTG TTG GAA CTG AAC AAG	1440
Arg Val Asn Ala Ala Asp Ile Glu Asn Arg Val Leu Glu Leu Asn Lys	
465 470 475 480	
AAG CAG GAG TCC GAG GAT ACA GCC AAG GCT GGC TTC TGG GAG GAG TTT	1488
Lys Gln Glu Ser Glu Asp Thr Ala Lys Ala Gly Phe Trp Glu Glu Phe	
485 490 495	
GAG AGT TTG CAG AAG CAG GAG GTG AAG AAC TTG CAC CAG CGT CTG GAA	1536
Glu Ser Leu Gln Lys Gln Glu Val Lys Asn Leu His Gln Arg Leu Glu	
500 505 510	
GGG CAG CGG CCA GAG AAC AAG GGC AAG AAC CGC TAC AAG AAC ATT CTC	1584
Gly Gln Arg Pro Glu Asn Lys Gly Lys Asn Arg Tyr Lys Asn Ile Leu	
515 520 525	
CCC TTT GAC CAC AGC CGA GTG ATC CTG CAG GGA CGG GAC AGT AAC ATC	1632
Pro Phe Asp His Ser Arg Val Ile Leu Gln Gly Arg Asp Ser Asn Ile	
530 535 540	
CCC GGG TCC GAC TAC ATC AAT GCC AAC TAC ATC AAG AAC CAG CTG CTA	1680
Pro Gly Ser Asp Tyr Ile Asn Ala Asn Tyr Ile Lys Asn Gln Leu Leu	
545 550 555 560	
GGC CCT GAT GAG AAC GCT AAG ACC TAC ATC GCC AGC CAG GGC TGT CTG	1728
Gly Pro Asp Glu Asn Ala Lys Thr Tyr Ile Ala Ser Gln Gly Cys Leu	
565 570 575	

GAG GCC ACG GTC AAT GAC TTC TGG CAG ATG GCG TGG CAG GAG AAC AGC Glu Ala Thr Val Asn Asp Phe Trp Gln Met Ala Trp Gln Glu Asn Ser 580 585 590	1776
CGT GTC ATC GTC ATG ACC ACC CGA GAG GTG GAG AAA GGC CGG AAC AAA Arg Val Ile Val Met Thr Thr Arg Glu Val Glu Lys Gly Arg Asn Lys 595 600 605	1824
TGC GTC CCA TAC TGG CCC GAG GTG GGC ATG CAG CGT GCT TAT GGG CCC Cys Val Pro Tyr Trp Pro Glu Val Gly Met Gln Arg Ala Tyr Gly Pro 610 615 620	1872
TAC TCT GTG ACC AAC TGC GGG GAG CAT GAC ACA ACC GAA TAC AAA CTC Tyr Ser Val Thr Asn Cys Gly Glu His Asp Thr Thr Glu Tyr Lys Leu 625 630 635 640	1920
CGT ACC TTA CAG GTC TCC CCG CTG GAC AAT GGA GAC CTG ATT CGG GAG Arg Thr Leu Gln Val Ser Pro Leu Asp Asn Gly Asp Leu Ile Arg Glu 645 650 655	1968
ATC TGG CAT TAC CAG TAC CTG AGC TGG CCC GAC CAT GGG GTC CCC AGT Ile Trp His Tyr Gln Tyr Leu Ser Trp Pro Asp His Gly Val Pro Ser 660 665 670	2016
GAG CCT GGG GGT GTC CTC AGC TTC CTG GAC CAG ATC AAC CAG CGG CAG Glu Pro Gly Gly Val Leu Ser Phe Leu Asp Gln Ile Asn Gln Arg Gln 675 680 685	2064
GAA AGT CTG CCT CAC GCA GGG CCC ATC ATC GTG CAC TGC AGC GCC GGC Glu Ser Leu Pro His Ala Gly Pro Ile Ile Val His Cys Ser Ala Gly 690 695 700	2112
ATC GGC CGC ACA GGC ACC ATC ATT GTC ATC GAC ATG CTC ATG GAG AAC Ile Gly Arg Thr Gly Thr Ile Ile Val Ile Asp Met Leu Met Glu Asn 705 710 715 720	2160
ATC TCC ACC AAG GGC CTG GAC TGT GAC ATT GAC ATC CAG AAG ACC ATC Ile Ser Thr Lys Gly Leu Asp Cys Asp Ile Asp Ile Gln Lys Thr Ile 725 730 735	2208
CAG ATG GTG CGG GCG CAG CGC TCG GGC ATG GTG CAG ACG GAG GCG CAG Gln Met Val Arg Ala Gln Arg Ser Gly Met Val Gln Thr Glu Ala Gln 740 745 750	2256
TAC AAG TTC ATC TAC GTG GCC ATC GCC CAG TTC ATT GAA ACC ACT AAG Tyr Lys Phe Ile Tyr Val Ala Ile Ala Gln Phe Ile Glu Thr Thr Lys 755 760 765	2304
AAG AAG CTG GAG GTC CTG CAG TCG CAG AAG GGC CAG GAG TCG GAG TAC Lys Lys Leu Glu Val Leu Gln Ser Gln Lys Gly Gln Glu Ser Glu Tyr 770 775 780	2352
GGG AAC ATC ACC TAT CCC CCA GCC ATG AAG AAT GCC CAT GCC AAG GCC Gly Asn Ile Thr Tyr Pro Pro Ala Met Lys Asn Ala His Ala Lys Ala 785 790 795 800	2400
TCC CGC ACC TCG TCC AAA CAC AAG GAG GAT GTG TAT GAG AAC CTG CAC	2448

Ser Arg Thr Ser Ser Lys His Lys Glu Asp Val Tyr Glu Asn Leu His
805 810 815

ACT AAG AAC AAG AGG GAG GAG AAA GTG AAG AAG CAG CGG TCA GCA GAC 2496
Thr Lys Asn Lys Arg Glu Glu Lys Val Lys Lys Gln Arg Ser Ala Asp
820 825 830

AAG GAG AAG AGC AAG GGT TCC CTC AAG AGG AAG TGA 2532
Lys Glu Lys Ser Lys Gly Ser Leu Lys Arg Lys
835 840

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 843 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1 5 10 15
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
20 25 30
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
35 40 45
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
50 55 60
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65 70 75 80
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
85 90 95
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
100 105 110
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
115 120 125
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
130 135 140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145 150 155 160
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
165 170 175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
180 185 190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
195 200 205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
210 215 220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
225 230 235 240
Gly Leu Arg Ser Arg Glu Met Leu Ser Arg Gly Trp Phe His Arg Asp
245 250 255
Leu Ser Gly Leu Asp Ala Glu Thr Leu Leu Lys Gly Arg Gly Val His

	260		265		270
Gly Ser Phe Leu Ala Arg Pro Ser Arg Lys Asn Gln Gly Asp Phe Ser					
	275		280		285
Leu Ser Val Arg Val Gly Asp Gln Val Thr His Ile Arg Ile Gln Asn					
	290		295		300
Ser Gly Asp Phe Tyr Asp Leu Tyr Gly Gly Glu Lys Phe Ala Thr Leu					
305		310		315	320
Thr Glu Leu Val Glu Tyr Tyr Thr Gln Gln Gln Gly Val Leu Gln Asp					
	325		330		335
Arg Asp Gly Thr Ile Ile His Leu Lys Tyr Pro Leu Asn Cys Ser Asp					
	340		345		350
Pro Thr Ser Glu Arg Trp Tyr His Gly His Met Ser Gly Gln Ala					
	355		360		365
Glu Thr Leu Leu Gln Ala Lys Gly Glu Pro Trp Thr Phe Leu Val Arg					
	370		375		380
Glu Ser Leu Ser Gln Pro Gly Asp Phe Val Leu Ser Val Leu Ser Asp					
385		390		395	400
Gln Pro Lys Ala Gly Pro Gly Ser Pro Leu Arg Val Thr His Ile Lys					
	405		410		415
Val Met Cys Glu Gly Gly Arg Tyr Thr Val Gly Gly Leu Glu Thr Phe					
	420		425		430
Asp Ser Leu Thr Asp Leu Val Glu His Phe Lys Lys Thr Gly Ile Glu					
	435		440		445
Glu Ala Ser Gly Ala Phe Val Tyr Leu Arg Gln Pro Tyr Tyr Ala Thr					
	450		455		460
Arg Val Asn Ala Ala Asp Ile Glu Asn Arg Val Leu Glu Leu Asn Lys					
465		470		475	480
Lys Gln Glu Ser Glu Asp Thr Ala Lys Ala Gly Phe Trp Glu Glu Phe					
	485		490		495
Glu Ser Leu Gln Lys Gln Glu Val Lys Asn Leu His Gln Arg Leu Glu					
	500		505		510
Gly Gln Arg Pro Glu Asn Lys Gly Lys Asn Arg Tyr Lys Asn Ile Leu					
	515		520		525
Pro Phe Asp His Ser Arg Val Ile Leu Gln Gly Arg Asp Ser Asn Ile					
	530		535		540
Pro Gly Ser Asp Tyr Ile Asn Ala Asn Tyr Ile Lys Asn Gln Leu Leu					
545		550		555	560
Gly Pro Asp Glu Asn Ala Lys Thr Tyr Ile Ala Ser Gln Gly Cys Leu					
	565		570		575
Glu Ala Thr Val Asn Asp Phe Trp Gln Met Ala Trp Gln Glu Asn Ser					
	580		585		590
Arg Val Ile Val Met Thr Thr Arg Glu Val Glu Lys Gly Arg Asn Lys					
	595		600		605
Cys Val Pro Tyr Trp Pro Glu Val Gly Met Gln Arg Ala Tyr Gly Pro					
	610		615		620
Tyr Ser Val Thr Asn Cys Gly Glu His Asp Thr Thr Glu Tyr Lys Leu					
625		630		635	640
Arg Thr Leu Gln Val Ser Pro Leu Asp Asn Gly Asp Leu Ile Arg Glu					
	645		650		655
Ile Trp His Tyr Gln Tyr Leu Ser Trp Pro Asp His Gly Val Pro Ser					
	660		665		670
Glu Pro Gly Gly Val Leu Ser Phe Leu Asp Gln Ile Asn Gln Arg Gln					
	675		680		685
Glu Ser Leu Pro His Ala Gly Pro Ile Ile Val His Cys Ser Ala Gly					
	690		695		700
Ile Gly Arg Thr Gly Thr Ile Ile Val Ile Asp Met Leu Met Glu Asn					
705		710		715	720
Ile Ser Thr Lys Gly Leu Asp Cys Asp Ile Asp Ile Gln Lys Thr Ile					

	725		730		735
Gln Met Val Arg Ala Gln Arg Ser Gly Met Val Gln Thr Glu Ala Gln					
	740		745		750
Tyr Lys Phe Ile Tyr Val Ala Ile Ala Gln Phe Ile Glu Thr Thr Lys					
	755		760		765
Lys Lys Leu Glu Val Leu Gln Ser Gln Lys Gly Gln Glu Ser Glu Tyr					
	770		775		780
Gly Asn Ile Thr Tyr Pro Pro Ala Met Lys Asn Ala His Ala Lys Ala					
785		790		795	800
Ser Arg Thr Ser Ser Lys His Lys Glu Asp Val Tyr Glu Asn Leu His					
	805		810		815
Thr Lys Asn Lys Arg Glu Glu Lys Val Lys Lys Gln Arg Ser Ala Asp					
	820		825		830
Lys Glu Lys Ser Lys Gly Ser Leu Lys Arg Lys					
	835		840		

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2562 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2559
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

ATG CTG TCC CGT GGG TGG TTT CAC CGA GAC CTC AGT GGG CTG GAT GCA	48
Met Leu Ser Arg Gly Trp Phe His Arg Asp Leu Ser Gly Leu Asp Ala	
1 5 10 15	
GAG ACC CTG CTC AAG GGC CGA GGT GTC CAC GGT AGC TTC CTG GCT CGG	96
Glu Thr Leu Leu Lys Gly Arg Gly Val His Gly Ser Phe Leu Ala Arg	
20 25 30	
CCC AGT CGC AAG AAC CAG GGT GAC TTC TCG CTC TCC GTC AGG GTG GGG	144
Pro Ser Arg Lys Asn Gln Gly Asp Phe Ser Leu Ser Val Arg Val Gly	
35 40 45	
GAT CAG GTG ACC CAT ATT CGG ATC CAG AAC TCA GGG GAT TTC TAT GAC	192
Asp Gln Val Thr His Ile Arg Ile Gln Asn Ser Gly Asp Phe Tyr Asp	
50 55 60	
CTG TAT GGA GGG GAG AAG TTT GCG ACT CTG ACA GAG CTG GTG GAG TAC	240
Leu Tyr Gly Gly Glu Lys Phe Ala Thr Leu Thr Glu Leu Val Glu Tyr	
65 70 75 80	
TAC ACT CAG CAG CAG GGT GTC CTG CAG GAC CGC GAC GGC ACC ATC ATC	288
Tyr Thr Gln Gln Gln Gly Val Leu Gln Asp Arg Asp Gly Thr Ile Ile	
85 90 95	
CAC CTC AAG TAC CCG CTG AAC TGC TCC GAT CCC ACT AGT GAG AGG TGG	336

His	Leu	Lys	Tyr	Pro	Leu	Asn	Cys	Ser	Asp	Pro	Thr	Ser	Glu	Arg	Trp	
			100					105					110			
TAC	CAT	GGC	CAC	ATG	TCT	GGC	GGG	CAG	GCA	GAG	ACG	CTG	CTG	CAG	GCC	384
Tyr	His	Gly	His	Met	Ser	Gly	Gly	Gln	Ala	Glu	Thr	Leu	Leu	Gln	Ala	
		115				120				125						
AAG	GGC	GAG	CCC	TGG	ACG	TTT	CTT	GTG	CGT	GAG	AGC	CTC	AGC	CAG	CCT	432
Lys	Gly	Glu	Pro	Trp	Thr	Phe	Leu	Val	Arg	Glu	Ser	Leu	Ser	Gln	Pro	
	130					135				140						
GGA	GAC	TTC	GTG	CTT	TCT	GTG	CTC	AGT	GAC	CAG	CCC	AAG	GCT	GGC	CCA	480
Gly	Asp	Phe	Val	Leu	Ser	Val	Leu	Ser	Asp	Gln	Pro	Lys	Ala	Gly	Pro	
145					150				155					160		
GGC	TCC	CCG	CTC	AGG	GTC	ACC	CAC	ATC	AAG	GTC	ATG	TGC	GAG	GGT	GGA	528
Gly	Ser	Pro	Leu	Arg	Val	Thr	His	Ile	Lys	Val	Met	Cys	Glu	Gly	Gly	
			165					170					175			
CGC	TAC	ACA	GTG	GGT	GGT	TTG	GAG	ACC	TTC	GAC	AGC	CTC	ACG	GAC	CTG	576
Arg	Tyr	Thr	Val	Gly	Gly	Leu	Glu	Thr	Phe	Asp	Ser	Leu	Thr	Asp	Leu	
		180					185					190				
GTA	GAG	CAT	TTC	AAG	AAG	ACG	GGG	ATT	GAG	GAG	GCC	TCA	GGC	GCC	TTT	624
Val	Glu	His	Phe	Lys	Lys	Thr	Gly	Ile	Glu	Glu	Ala	Ser	Gly	Ala	Phe	
	195					200					205					
GTC	TAC	CTG	CGG	CAG	CCG	TAC	TAT	GCC	ACG	AGG	GTG	AAT	GCG	GCT	GAC	672
Val	Tyr	Leu	Arg	Gln	Pro	Tyr	Tyr	Ala	Thr	Arg	Val	Asn	Ala	Ala	Asp	
	210					215					220					
ATT	GAG	AAC	CGA	GTG	TTG	GAA	CTG	AAC	AAG	AAG	CAG	GAG	TCC	GAG	GAT	720
Ile	Glu	Asn	Arg	Val	Leu	Glu	Leu	Asn	Lys	Lys	Gln	Glu	Ser	Glu	Asp	
225					230					235				240		
ACA	GCC	AAG	GCT	GGC	TTC	TGG	GAG	GAG	TTT	GAG	ACT	TTG	CAG	AAG	CAG	768
Thr	Ala	Lys	Ala	Gly	Phe	Trp	Glu	Glu	Phe	Glu	Ser	Leu	Gln	Lys	Gln	
			245					250					255			
GAG	GTG	AAG	AAC	TTG	CAC	CAG	CGT	CTG	GAA	GGG	CAG	CGG	CCA	GAG	AAC	816
Glu	Val	Lys	Asn	Leu	His	Gln	Arg	Leu	Glu	Gly	Gln	Arg	Pro	Glu	Asn	
		260						265					270			
AAG	GGC	AAG	AAC	CGC	TAC	AAG	AAC	ATT	CTC	CCC	TTT	GAC	CAC	AGC	CGA	864
Lys	Gly	Lys	Asn	Arg	Tyr	Lys	Asn	Ile	Leu	Pro	Phe	Asp	His	Ser	Arg	
	275						280					285				
GTG	ATC	CTG	CAG	GGA	CGG	GAC	AGT	AAC	ATC	CCC	GGG	TCC	GAC	TAC	ATC	912
Val	Ile	Leu	Gln	Gly	Arg	Asp	Ser	Asn	Ile	Pro	Gly	Ser	Asp	Tyr	Ile	
	290					295					300					
AAT	GCC	AAC	TAC	ATC	AAG	AAC	CAG	CTG	CTA	GGC	CCT	GAT	GAG	AAC	GCT	960
Asn	Ala	Asn	Tyr	Ile	Lys	Asn	Gln	Leu	Leu	Gly	Pro	Asp	Glu	Asn	Ala	
305					310					315				320		
AAG	ACC	TAC	ATC	GCC	AGC	CAG	GGC	TGT	CTG	GAG	GCC	ACG	GTC	AAT	GAC	1008
Lys	Thr	Tyr	Ile	Ala	Ser	Gln	Gly	Cys	Leu	Glu	Ala	Thr	Val	Asn	Asp	
			325					330					335			

TTC TGG CAG ATG GCG TGG CAG GAG AAC AGC CGT GTC ATC GTC ATG ACC Phe Trp Gln Met Ala Trp Gln Glu Asn Ser Arg Val Ile Val Met Thr 340 345 350	1056
ACC CGA GAG GTG GAG AAA GGC CGG AAC AAA TGC GTC CCA TAC TGG CCC Thr Arg Glu Val Glu Lys Gly Arg Asn Lys Cys Val Pro Tyr Trp Pro 355 360 365	1104
GAG GTG GGC ATG CAG CGT GCT TAT GGG CCC TAC TCT GTG ACC AAC TGC Glu Val Gly Met Gln Arg Ala Tyr Gly Pro Tyr Ser Val Thr Asn Cys 370 375 380	1152
GGG GAG CAT GAC ACA ACC GAA TAC AAA CTC CGT ACC TTA CAG GTC TCC Gly Glu His Asp Thr Glu Tyr Lys Leu Arg Thr Leu Gln Val Ser 385 390 395 400	1200
CCG CTG GAC AAT GGA GAC CTG ATT CGG GAG ATC TGG CAT TAC CAG TAC Pro Leu Asp Asn Gly Asp Leu Ile Arg Glu Ile Trp His Tyr Gln Tyr 405 410 415	1248
CTG AGC TGG CCC GAC CAT GGG GTC CCC AGT GAG CCT GGG GGT GTC CTC Leu Ser Trp Pro Asp His Gly Val Pro Ser Glu Pro Gly Gly Val Leu 420 425 430	1296
AGC TTC CTG GAC CAG ATC AAC CAG CGG CAG GAA AGT CTG CCT CAC GCA Ser Phe Leu Asp Gln Ile Asn Gln Arg Gln Glu Ser Leu Pro His Ala 435 440 445	1344
GGG CCC ATC ATC GTG CAC TGC AGC GCC GGC ATC GGC CGC ACA GGC ACC Gly Pro Ile Ile Val His Cys Ser Ala Gly Ile Gly Arg Thr Gly Thr 450 455 460	1392
ATC ATT GTC ATC GAC ATG CTC ATG GAG AAC ATC TCC ACC AAG GGC CTG Ile Ile Val Ile Asp Met Leu Met Glu Asn Ile Ser Thr Lys Gly Leu 465 470 475 480	1440
GAC TGT GAC ATT GAC ATC CAG AAG ACC ATC CAG ATG GTG CGG GCG CAG Asp Cys Asp Ile Asp Ile Gln Lys Thr Ile Gln Met Val Arg Ala Gln 485 490 495	1488
CGC TCG GGC ATG GTG CAG ACG GAG GCG CAG TAC AAG TTC ATC TAC GTG Arg Ser Gly Met Val Gln Thr Glu Ala Gln Tyr Lys Phe Ile Tyr Val 500 505 510	1536
GCC ATC GCC CAG TTC ATT GAA ACC ACT AAG AAG AAG CTG GAG GTC CTG Ala Ile Ala Gln Phe Ile Glu Thr Thr Lys Lys Lys Leu Glu Val Leu 515 520 525	1584
CAG TCG CAG AAG GGC CAG GAG TCG GAG TAC GGG AAC ATC ACC TAT CCC Gln Ser Gln Lys Gly Gln Glu Ser Glu Tyr Gly Asn Ile Thr Tyr Pro 530 535 540	1632
CCA GCC ATG AAG AAT GCC CAT GCC AAG GCC TCC CGC ACC TCG TCC AAA Pro Ala Met Lys Asn Ala His Ala Lys Ala Ser Arg Thr Ser Ser Lys 545 550 555 560	1680
CAC AAG GAG GAT GTG TAT GAG AAC CTG CAC ACT AAG AAC AAG AGG GAG	1728

His Lys Glu Asp Val Tyr Glu Asn Leu His Thr Lys Asn Lys Arg Glu	
565 570 575	
GAG AAA GTG AAG AAG CAG CGG TCA GCA GAC AAG GAG AAG AGC AAG GGT	1776
Glu Lys Val Lys Lys Gln Arg Ser Ala Asp Lys Glu Lys Ser Lys Gly	
580 585 590	
TCC CTC AAG AGG AAG CGA ATT CTG CAG TCG ACG GTA CCG CGG GCC CGG	1824
Ser Leu Lys Arg Lys Arg Ile Leu Gln Ser Thr Val Pro Arg Ala Arg	
595 600 605	
GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC	1872
Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr	
610 615 620	
GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC	1920
Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His	
625 630 635 640	
AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG	1968
Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys	
645 650 655	
CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG	2016
Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp	
660 665 670	
CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC	2064
Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg	
675 680 685	
TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC	2112
Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro	
690 695 700	
GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC	2160
Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn	
705 710 715 720	
TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC	2208
Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn	
725 730 735	
CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG	2256
Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu	
740 745 750	
GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG	2304
Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met	
755 760 765	
GTC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC	2352
Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His	
770 775 780	
AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC	2400
Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn	
785 790 795 800	

ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG 2448
 Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu
 805 810 815

AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC 2496
 Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His
 820 825 830

ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG 2544
 Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met
 835 840 845

GAC GAG CTG TAC AAG TAA 2562
 Asp Glu Leu Tyr Lys
 850

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 853 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

Met Leu Ser Arg Gly Trp Phe His Arg Asp Leu Ser Gly Leu Asp Ala
 1 5 10 15
 Glu Thr Leu Leu Lys Gly Arg Gly Val His Gly Ser Phe Leu Ala Arg
 20 25 30
 Pro Ser Arg Lys Asn Gln Gly Asp Phe Ser Leu Ser Val Arg Val Gly
 35 40 45
 Asp Gln Val Thr His Ile Arg Ile Gln Asn Ser Gly Asp Phe Tyr Asp
 50 55 60
 Leu Tyr Gly Gly Glu Lys Phe Ala Thr Leu Thr Glu Leu Val Glu Tyr
 65 70 75 80
 Tyr Thr Gln Gln Gln Gly Val Leu Gln Asp Arg Asp Gly Thr Ile Ile
 85 90 95
 His Leu Lys Tyr Pro Leu Asn Cys Ser Asp Pro Thr Ser Glu Arg Trp
 100 105 110
 Tyr His Gly His Met Ser Gly Gly Gln Ala Glu Thr Leu Leu Gln Ala
 115 120 125
 Lys Gly Glu Pro Trp Thr Phe Leu Val Arg Glu Ser Leu Ser Gln Pro
 130 135 140
 Gly Asp Phe Val Leu Ser Val Leu Ser Asp Gln Pro Lys Ala Gly Pro
 145 150 155 160
 Gly Ser Pro Leu Arg Val Thr His Ile Lys Val Met Cys Glu Gly Gly
 165 170 175
 Arg Tyr Thr Val Gly Gly Leu Glu Thr Phe Asp Ser Leu Thr Asp Leu
 180 185 190
 Val Glu His Phe Lys Lys Thr Gly Ile Glu Glu Ala Ser Gly Ala Phe
 195 200 205
 Val Tyr Leu Arg Gln Pro Tyr Tyr Ala Thr Arg Val Asn Ala Ala Asp

210	215	220
Ile Glu Asn Arg Val Leu Glu Leu Asn Lys Lys Gln Glu Ser Glu Asp		
225	230	235
Thr Ala Lys Ala Gly Phe Trp Glu Glu Phe Glu Ser Leu Gln Lys Gln		240
	245	250
Glu Val Lys Asn Leu His Gln Arg Leu Glu Gly Gln Arg Pro Glu Asn		255
	260	265
Lys Gly Lys Asn Arg Tyr Lys Asn Ile Leu Pro Phe Asp His Ser Arg		270
	275	280
Val Ile Leu Gln Gly Arg Asp Ser Asn Ile Pro Gly Ser Asp Tyr Ile		285
	290	295
Asn Ala Asn Tyr Ile Lys Asn Gln Leu Leu Gly Pro Asp Glu Asn Ala		300
305	310	315
Lys Thr Tyr Ile Ala Ser Gln Gly Cys Leu Glu Ala Thr Val Asn Asp		320
	325	330
Phe Trp Gln Met Ala Trp Gln Glu Asn Ser Arg Val Ile Val Met Thr		335
	340	345
Thr Arg Glu Val Glu Lys Gly Arg Asn Lys Cys Val Pro Tyr Trp Pro		350
	355	360
Glu Val Gly Met Gln Arg Ala Tyr Gly Pro Tyr Ser Val Thr Asn Cys		365
	370	375
Gly Glu His Asp Thr Thr Glu Tyr Lys Leu Arg Thr Leu Gln Val Ser		380
385	390	395
Pro Leu Asp Asn Gly Asp Leu Ile Arg Glu Ile Trp His Tyr Gln Tyr		400
	405	410
Leu Ser Trp Pro Asp His Gly Val Pro Ser Glu Pro Gly Gly Val Leu		415
	420	425
Ser Phe Leu Asp Gln Ile Asn Gln Arg Gln Glu Ser Leu Pro His Ala		430
	435	440
Gly Pro Ile Ile Val His Cys Ser Ala Gly Ile Gly Arg Thr Gly Thr		445
	450	455
Ile Ile Val Ile Asp Met Leu Met Glu Asn Ile Ser Thr Lys Gly Leu		460
465	470	475
Asp Cys Asp Ile Asp Ile Gln Lys Thr Ile Gln Met Val Arg Ala Gln		480
	485	490
Arg Ser Gly Met Val Gln Thr Glu Ala Gln Tyr Lys Phe Ile Tyr Val		495
	500	505
Ala Ile Ala Gln Phe Ile Glu Thr Thr Lys Lys Lys Leu Glu Val Leu		510
	515	520
Gln Ser Gln Lys Gly Gln Glu Ser Glu Tyr Gly Asn Ile Thr Tyr Pro		525
	530	535
Pro Ala Met Lys Asn Ala His Ala Lys Ala Ser Arg Thr Ser Ser Lys		540
545	550	555
His Lys Glu Asp Val Tyr Glu Asn Leu His Thr Lys Asn Lys Arg Glu		560
	565	570
Glu Lys Val Lys Lys Gln Arg Ser Ala Asp Lys Glu Lys Ser Lys Gly		575
	580	585
Ser Leu Lys Arg Lys Arg Ile Leu Gln Ser Thr Val Pro Arg Ala Arg		590
	595	600
Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr		605
	610	615
Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His		620
625	630	635
Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys		640
	645	650
Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp		655
	660	665
Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg		670

675	680	685
Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro		
690	695	700
Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn		
705	710	715
Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn		
725	730	735
Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu		
740	745	750
Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met		
755	760	765
Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His		
770	775	780
Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn		
785	790	795
Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu		
805	810	815
Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His		
820	825	830
Met Val Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met		
835	840	845
Asp Glu Leu Tyr Lys		
850		

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2994 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2991
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240

Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	
65					70					75					80	
CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	288
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	
			85						90					95		
CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	336
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	
			100					105					110			
GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	CTG	AAG	GGC	384
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	
			115				120					125				
ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	CTG	GAG	TAC	432
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	
			130			135					140					
AAC	TAC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	480
Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	
			145			150				155					160	
GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	528
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	
				165				170						175		
GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	576
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	
			180					185					190			
CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	TCC	GCC	CTG	624
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	
			195				200					205				
AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	CTG	GAG	TTC	672
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	
			210				215					220				
GTG	ACC	GCC	GCC	GGG	ATC	ACT	CTC	GGC	ATG	GAC	GAG	CTG	TAC	AAG	TCC	720
Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Ser	
			225			230				235					240	
GGA	CTC	AGA	TCT	CGA	GCT	CAA	GCT	TCG	AAT	TCG	ACC	ATG	GAG	CGG	CCC	768
Gly	Leu	Arg	Ser	Arg	Ala	Gln	Ala	Ser	Asn	Ser	Thr	Met	Glu	Arg	Pro	
				245					250					255		
CCG	GGG	CTG	CGG	CCG	GGC	GCG	GGC	GGG	CCC	TGG	GAG	ATG	CGG	GAG	CGG	816
Pro	Gly	Leu	Arg	Pro	Gly	Ala	Gly	Gly	Pro	Trp	Glu	Met	Arg	Glu	Arg	
			260					265					270			
CTG	GGC	ACC	GGC	GGC	TTC	GGG	AAC	GTC	TGT	CTG	TAC	CAG	CAT	CGG	GAA	864
Leu	Gly	Thr	Gly	Gly	Phe	Gly	Asn	Val	Cys	Leu	Tyr	Gln	His	Arg	Glu	
			275				280					285				
CTT	GAT	CTC	AAA	ATA	GCA	ATT	AAG	TCT	TGT	CGC	CTA	GAG	CTA	AGT	ACC	912
Leu	Asp	Leu	Lys	Ile	Ala	Ile	Lys	Ser	Cys	Arg	Leu	Glu	Leu	Ser	Thr	
			290				295					300				

AAA AAC AGA GAA CGA TGG TGC CAT GAA ATC CAG ATT ATG AAG AAG TTG	960
Lys Asn Arg Glu Arg Trp Cys His Glu Ile Gln Ile Met Lys Lys Leu	
305 310 315 320	
AAC CAT GCC AAT GTT GTA AAG GCC TGT GAT GTT CCT GAA GAA TTG AAT	1008
Asn His Ala Asn Val Val Lys Ala Cys Asp Val Pro Glu Glu Leu Asn	
325 330 335	
ATT TTG ATT CAT GAT GTG CCT CTT CTA GCA ATG GAA TAC TGT TCT GGA	1056
Ile Leu Ile His Asp Val Pro Leu Leu Ala Met Glu Tyr Cys Ser Gly	
340 345 350	
GGA GAT CTC CGA AAG CTG CTC AAC AAA CCA GAA AAT TGT TGT GGA CTT	1104
Gly Asp Leu Arg Lys Leu Leu Asn Lys Pro Glu Asn Cys Cys Gly Leu	
355 360 365	
AAA GAA AGC CAG ATA CTT TCT TTA CTA AGT GAT ATA GGG TCT GGG ATT	1152
Lys Glu Ser Gln Ile Leu Ser Leu Leu Ser Asp Ile Gly Ser Gly Ile	
370 375 380	
CGA TAT TTG CAT GAA AAC AAA ATT ATA CAT CGA GAT CTA AAA CCT GAA	1200
Arg Tyr Leu His Glu Asn Lys Ile Ile His Arg Asp Leu Lys Pro Glu	
385 390 395 400	
AAC ATA GTT CTT CAG GAT GTT GGT GGA AAG ATA ATA CAT AAA ATA ATT	1248
Asn Ile Val Leu Gln Asp Val Gly Gly Lys Ile Ile His Lys Ile Ile	
405 410 415	
GAT CTG GGA TAT GCC AAA GAT GTT GAT CAA GGA AGT CTG TGT ACA TCT	1296
Asp Leu Gly Tyr Ala Lys Asp Val Asp Gln Gly Ser Leu Cys Thr Ser	
420 425 430	
TTT GTG GGA ACA CTG CAG TAT CTG GCC CCA GAG CTC TTT GAG AAT AAG	1344
Phe Val Gly Thr Leu Gln Tyr Leu Ala Pro Glu Leu Phe Glu Asn Lys	
435 440 445	
CCT TAC ACA GCC ACT GTT GAT TAT TGG AGC TTT GGG ACC ATG GTA TTT	1392
Pro Tyr Thr Ala Thr Val Asp Tyr Trp Ser Phe Gly Thr Met Val Phe	
450 455 460	
GAA TGT ATT GCT GGA TAT AGG CCT TTT TTG CAT CAT CTG CAG CCA TTT	1440
Glu Cys Ile Ala Gly Tyr Arg Pro Phe Leu His His Leu Gln Pro Phe	
465 470 475 480	
ACC TGG CAT GAG AAG ATT AAG AAG AAG GAT CCA AAG TGT ATA TTT GCA	1488
Thr Trp His Glu Lys Ile Lys Lys Lys Asp Pro Lys Cys Ile Phe Ala	
485 490 495	
TGT GAA GAG ATG TCA GGA GAA GTT CGG TTT AGT AGC CAT TTA CCT CAA	1536
Cys Glu Glu Met Ser Gly Glu Val Arg Phe Ser Ser His Leu Pro Gln	
500 505 510	
CCA AAT AGC CTT TGT AGT TTA ATA GTA GAA CCC ATG GAA AAC TGG CTA	1584
Pro Asn Ser Leu Cys Ser Leu Ile Val Glu Pro Met Glu Asn Trp Leu	
515 520 525	
CAG TTG ATG TTG AAT TGG GAC CCT CAG CAG AGA GGA GGA CCT GTT GAC	1632

Gln Leu Met Leu Asn Trp Asp Pro Gln Gln Arg Gly Gly Pro Val Asp 530	535	540	
CTT ACT TTG AAG CAG CCA AGA TGT TTT GTA TTA ATG GAT CAC ATT TTG Leu Thr Leu Lys Gln Pro Arg Cys Phe Val Leu Met Asp His Ile Leu 545	550	555	1680 560
AAT TTG AAG ATA GTA CAC ATC CTA AAT ATG ACT TCT GCA AAG ATA ATT Asn Leu Lys Ile Val His Ile Leu Asn Met Thr Ser Ala Lys Ile Ile 565	570	575	1728
TCT TTT CTG TTA CCA CCT GAT GAA AGT CTT CAT TCA CTA CAG TCT CGT Ser Phe Leu Leu Pro Pro Asp Glu Ser Leu His Ser Leu Gln Ser Arg 580	585	590	1776
ATT GAG CGT GAA ACT GGA ATA AAT ACT GGT TCT CAA GAA CTT CTT TCA Ile Glu Arg Glu Thr Gly Ile Asn Thr Gly Ser Gln Glu Leu Leu Ser 595	600	605	1824
GAG ACA GGA ATT TCT CTG GAT CCT CGG AAA CCA GCC TCT CAA TGT GTT Glu Thr Gly Ile Ser Leu Asp Pro Arg Lys Pro Ala Ser Gln Cys Val 610	615	620	1872
CTA GAT GGA GTT AGA GGC TGT GAT AGC TAT ATG GTT TAT TTG TTT GAT Leu Asp Gly Val Arg Gly Cys Asp Ser Tyr Met Val Tyr Leu Phe Asp 625	630	635	1920 640
AAA AGT AAA ACT GTA TAT GAA GGG CCA TTT GCT TCC AGA AGT TTA TCT Lys Ser Lys Thr Val Tyr Glu Gly Pro Phe Ala Ser Arg Ser Leu Ser 645	650	655	1968
GAT TGT GTA AAT TAT ATT GTA CAG GAC AGC AAA ATA CAG CTT CCA ATT Asp Cys Val Asn Tyr Ile Val Gln Asp Ser Lys Ile Gln Leu Pro Ile 660	665	670	2016
ATA CAG CTG CGT AAA GTG TGG GCT GAA GCA GTG CAC TAT GTG TCT GGA Ile Gln Leu Arg Lys Val Trp Ala Glu Ala Val His Tyr Val Ser Gly 675	680	685	2064
CTA AAA GAA GAC TAT AGC AGG CTC TTT CAG GGA CAA AGG GCA GCA ATG Leu Lys Glu Asp Tyr Ser Arg Leu Phe Gln Gly Gln Arg Ala Ala Met 690	695	700	2112
TTA AGT CTT CTT AGA TAT AAT GCT AAC TTA ACA AAA ATG AAG AAC ACT Leu Ser Leu Leu Arg Tyr Asn Ala Asn Leu Thr Lys Met Lys Asn Thr 705	710	715	2160 720
TTG ATC TCA GCA TCA CAA CAA CTG AAA GCT AAA TTG GAG TTT TTT CAC Leu Ile Ser Ala Ser Gln Gln Leu Lys Ala Lys Leu Glu Phe Phe His 725	730	735	2208
AAA AGC ATT CAG CTT GAC TTG GAG AGA TAC AGC GAG CAG ATG ACG TAT Lys Ser Ile Gln Leu Asp Leu Glu Arg Tyr Ser Glu Gln Met Thr Tyr 740	745	750	2256
GGG ATA TCT TCA GAA AAA ATG CTA AAA GCA TGG AAA GAA ATG GAA GAA Gly Ile Ser Ser Glu Lys Met Leu Lys Ala Trp Lys Glu Met Glu Glu 755	760	765	2304

AAG GCC ATC CAC TAT GCT GAG GTT GGT GTC ATT GGA TAC CTG GAG GAT	2352
Lys Ala Ile His Tyr Ala Glu Val Gly Val Ile Gly Tyr Leu Glu Asp	
770 775 780	
CAG ATT ATG TCT TTG CAT GCT GAA ATC ATG GGG CTA CAG AAG AGC CCC	2400
Gln Ile Met Ser Leu His Ala Glu Ile Met Gly Leu Gln Lys Ser Pro	
785 790 795 800	
TAT GGA AGA CGT CAG GGA GAC TTG ATG GAA TCT CTG GAA CAG CGT GCC	2448
Tyr Gly Arg Arg Gln Gly Asp Leu Met Glu Ser Leu Glu Gln Arg Ala	
805 810 815	
ATT GAT CTA TAT AAG CAG TTA AAA CAC AGA CCT TCA GAT CAC TCC TAC	2496
Ile Asp Leu Tyr Lys Gln Leu Lys His Arg Pro Ser Asp His Ser Tyr	
820 825 830	
AGT GAC AGC ACA GAG ATG GTG AAA ATC ATT GTG CAC ACT GTG CAG AGT	2544
Ser Asp Ser Thr Glu Met Val Lys Ile Ile Val His Thr Val Gln Ser	
835 840 845	
CAG GAC CGT GTG CTC AAG GAG CTG TTT GGT CAT TTG AGC AAG TTG TTG	2592
Gln Asp Arg Val Leu Lys Glu Leu Phe Gly His Leu Ser Lys Leu Leu	
850 855 860	
GGC TGT AAG CAG AAG ATT ATT GAT CTA CTC CCT AAG GTG GAA GTG GCC	2640
Gly Cys Lys Gln Lys Ile Ile Asp Leu Leu Pro Lys Val Glu Val Ala	
865 870 875 880	
CTC AGT AAT ATC AAA GAA GCT GAC AAT ACT GTC ATG TTC ATG CAG GGA	2688
Leu Ser Asn Ile Lys Glu Ala Asp Asn Thr Val Met Phe Met Gln Gly	
885 890 895	
AAA AGG CAG AAA GAA ATA TGG CAT CTC CTT AAA ATT GCC TGT ACA CAG	2736
Lys Arg Gln Lys Glu Ile Trp His Leu Leu Lys Ile Ala Cys Thr Gln	
900 905 910	
AGT TCT GCC CGC TCT CTT GTA GGA TCC AGT CTA GAA GGT GCA GTA ACC	2784
Ser Ser Ala Arg Ser Leu Val Gly Ser Ser Leu Glu Gly Ala Val Thr	
915 920 925	
CCT CAG ACA TCA GCA TGG CTG CCC CCG ACT TCA GCA GAA CAT GAT CAT	2832
Pro Gln Thr Ser Ala Trp Leu Pro Pro Thr Ser Ala Glu His Asp His	
930 935 940	
TCT CTG TCA TGT GTG GTA ACT CCT CAA GAT GGG GAG ACT TCA GCA CAA	2880
Ser Leu Ser Cys Val Val Thr Pro Gln Asp Gly Glu Thr Ser Ala Gln	
945 950 955 960	
ATG ATA GAA GAA AAT TTG AAC TGC CTT GGC CAT TTA AGC ACT ATT ATT	2928
Met Ile Glu Glu Asn Leu Asn Cys Leu Gly His Leu Ser Thr Ile Ile	
965 970 975	
CAT GAG GCA AAT GAG GAA CAG GGC AAT AGT ATG ATG AAT CTT GAT TGG	2976
His Glu Ala Asn Glu Glu Gln Gly Asn Ser Met Met Asn Leu Asp Trp	
980 985 990	
AGT TGG TTA ACA GAA TGA	2994

Ser Trp Leu Thr Glu
995

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 997 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	1	5	10	15
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	20	25	30	
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	35	40	45	
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	50	55	60	
Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	65	70	75	80
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	85	90	95	
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	100	105	110	
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	115	120	125	
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	130	135	140	
Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	145	150	155	160
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	165	170	175	
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	180	185	190	
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	195	200	205	
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	210	215	220	
Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Ser	225	230	235	240
Gly	Leu	Arg	Ser	Arg	Ala	Gln	Ala	Ser	Asn	Ser	Thr	Met	Glu	Arg	Pro	245	250	255	
Pro	Gly	Leu	Arg	Pro	Gly	Ala	Gly	Gly	Pro	Trp	Glu	Met	Arg	Glu	Arg	260	265	270	
Leu	Gly	Thr	Gly	Gly	Phe	Gly	Asn	Val	Cys	Leu	Tyr	Gln	His	Arg	Glu	275	280	285	
Leu	Asp	Leu	Lys	Ile	Ala	Ile	Lys	Ser	Cys	Arg	Leu	Glu	Leu	Ser	Thr	290	295	300	
Lys	Asn	Arg	Glu	Arg	Trp	Cys	His	Glu	Ile	Gln	Ile	Met	Lys	Lys	Leu	305	310	315	320
Asn	His	Ala	Asn	Val	Val	Lys	Ala	Cys	Asp	Val	Pro	Glu	Glu	Leu	Asn				

[illegible]

785		790		795		800
Tyr Gly Arg Arg Gln Gly Asp Leu Met Glu Ser Leu Glu Gln Arg Ala						
	805			810		815
Ile Asp Leu Tyr Lys Gln Leu Lys His Arg Pro Ser Asp His Ser Tyr						
	820		825			830
Ser Asp Ser Thr Glu Met Val Lys Ile Ile Val His Thr Val Gln Ser						
	835		840			845
Gln Asp Arg Val Leu Lys Glu Leu Phe Gly His Leu Ser Lys Leu Leu						
	850		855			860
Gly Cys Lys Gln Lys Ile Ile Asp Leu Leu Pro Lys Val Glu Val Ala						
865		870		875		880
Leu Ser Asn Ile Lys Glu Ala Asp Asn Thr Val Met Phe Met Gln Gly						
	885		890			895
Lys Arg Gln Lys Glu Ile Trp His Leu Leu Lys Ile Ala Cys Thr Gln						
	900		905			910
Ser Ser Ala Arg Ser Leu Val Gly Ser Ser Leu Glu Gly Ala Val Thr						
	915		920			925
Pro Gln Thr Ser Ala Trp Leu Pro Pro Thr Ser Ala Glu His Asp His						
	930		935			940
Ser Leu Ser Cys Val Val Thr Pro Gln Asp Gly Glu Thr Ser Ala Gln						
945		950		955		960
Met Ile Glu Glu Asn Leu Asn Cys Leu Gly His Leu Ser Thr Ile Ile						
	965		970			975
His Glu Ala Asn Glu Glu Gln Gly Asn Ser Met Met Asn Leu Asp Trp						
	980		985			990
Ser Trp Leu Thr Glu						
	995					

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2991 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2988
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

ATG GAG CGG CCC CCG GGG CTG CGG CCG GGC GCG GGC GGG CCC TGG GAG	48
Met Glu Arg Pro Pro Gly Leu Arg Pro Gly Ala Gly Gly Pro Trp Glu	
1 5 10 15	
ATG CGG GAG CGG CTG GGC ACC GGC GGC TTC GCG AAC GTC TGT CTG TAC	96
Met Arg Glu Arg Leu Gly Thr Gly Gly Phe Gly Asn Val Cys Leu Tyr	
20 25 30	
CAG CAT CGG GAA CTT GAT CTC AAA ATA GCA ATT AAG TCT TGT CGC CTA	144
Gln His Arg Glu Leu Asp Leu Lys Ile Ala Ile Lys Ser Cys Arg Leu	
35 40 45	
GAG CTA AGT ACC AAA AAC AGA GAA CGA TGG TGC CAT GAA ATC CAG ATT	192

Glu Leu Ser Thr Lys Asn Arg Glu Arg Trp Cys His Glu Ile Gln Ile	
50 55 60	
ATG AAG AAG TTG AAC CAT GCC AAT GTT GTA AAG GCC TGT GAT GTT CCT	240
Met Lys Lys Leu Asn His Ala Asn Val Val Lys Ala Cys Asp Val Pro	
65 70 75 80	
GAA GAA TTG AAT ATT TTG ATT CAT GAT GTG CCT CTT CTA GCA ATG GAA	288
Glu Glu Leu Asn Ile Leu Ile His Asp Val Pro Leu Leu Ala Met Glu	
85 90 95	
TAC TGT TCT GGA GGA GAT CTC CGA AAG CTG CTC AAC AAA CCA GAA AAT	336
Tyr Cys Ser Gly Gly Asp Leu Arg Lys Leu Leu Asn Lys Pro Glu Asn	
100 105 110	
TGT TGT GGA CTT AAA GAA AGC CAG ATA CTT TCT TTA CTA AGT GAT ATA	384
Cys Cys Gly Leu Lys Glu Ser Gln Ile Leu Ser Leu Leu Ser Asp Ile	
115 120 125	
GGG TCT GGG ATT CGA TAT TTG CAT GAA AAC AAA ATT ATA CAT CGA GAT	432
Gly Ser Gly Ile Arg Tyr Leu His Glu Asn Lys Ile Ile His Arg Asp	
130 135 140	
CTA AAA CCT GAA AAC ATA GTT CTT CAG GAT GTT GGT GGA AAG ATA ATA	480
Leu Lys Pro Glu Asn Ile Val Leu Gln Asp Val Gly Gly Lys Ile Ile	
145 150 155 160	
CAT AAA ATA ATT GAT CTG GGA TAT GCC AAA GAT GTT GAT CAA GGA AGT	528
His Lys Ile Ile Asp Leu Gly Tyr Ala Lys Asp Val Asp Gln Gly Ser	
165 170 175	
CTG TGT ACA TCT TTT GTG GGA ACA CTG CAG TAT CTG GCC CCA GAG CTC	576
Leu Cys Thr Ser Phe Val Gly Thr Leu Gln Tyr Leu Ala Pro Glu Leu	
180 185 190	
TTT GAG AAT AAG CCT TAC ACA GCC ACT GTT GAT TAT TGG AGC TTT GGG	624
Phe Glu Asn Lys Pro Tyr Thr Ala Thr Val Asp Tyr Trp Ser Phe Gly	
195 200 205	
ACC ATG GTA TTT GAA TGT ATT GCT GGA TAT AGG CCT TTT TTG CAT CAT	672
Thr Met Val Phe Glu Cys Ile Ala Gly Tyr Arg Pro Phe Leu His His	
210 215 220	
CTG CAG CCA TTT ACC TGG CAT GAG AAG ATT AAG AAG AAG GAT CCA AAG	720
Leu Gln Pro Phe Thr Trp His Glu Lys Ile Lys Lys Lys Asp Pro Lys	
225 230 235 240	
TGT ATA TTT GCA TGT GAA GAG ATG TCA GGA GAA GTT CGG TTT AGT AGC	768
Cys Ile Phe Ala Cys Glu Glu Met Ser Gly Glu Val Arg Phe Ser Ser	
245 250 255	
CAT TTA CCT CAA CCA AAT AGC CTT TGT AGT TTA ATA GTA GAA CCC ATG	816
His Leu Pro Gln Pro Asn Ser Leu Cys Ser Leu Ile Val Glu Pro Met	
260 265 270	
GAA AAC TGG CTA CAG TTG ATG TTG AAT TGG GAC CCT CAG CAG AGA GGA	864
Glu Asn Trp Leu Gln Leu Met Leu Asn Trp Asp Pro Gln Gln Arg Gly	
275 280 285	

GGA CCT GTT GAC CTT ACT TTG AAG CAG CCA AGA TGT TTT GTA TTA ATG Gly Pro Val Asp Leu Thr Leu Lys Gln Pro Arg Cys Phe Val Leu Met 290 295 300	912
GAT CAC ATT TTG AAT TTG AAG ATA GTA CAC ATC CTA AAT ATG ACT TCT Asp His Ile Leu Asn Leu Lys Ile Val His Ile Leu Asn Met Thr Ser 305 310 315 320	960
GCA AAG ATA ATT TCT TTT CTG TTA CCA CCT GAT GAA AGT CTT CAT TCA Ala Lys Ile Ile Ser Phe Leu Leu Pro Pro Asp Glu Ser Leu His Ser 325 330 335	1008
CTA CAG TCT CGT ATT GAG CGT GAA ACT GGA ATA AAT ACT GGT TCT CAA Leu Gln Ser Arg Ile Glu Arg Glu Thr Gly Ile Asn Thr Gly Ser Gln 340 345 350	1056
GAA CTT CTT TCA GAG ACA GGA ATT TCT CTG GAT CCT CGG AAA CCA GCC Glu Leu Leu Ser Glu Thr Gly Ile Ser Leu Asp Pro Arg Lys Pro Ala 355 360 365	1104
TCT CAA TGT GTT CTA GAT GGA GTT AGA GGC TGT GAT AGC TAT ATG GTT Ser Gln Cys Val Leu Asp Gly Val Arg Gly Cys Asp Ser Tyr Met Val 370 375 380	1152
TAT TTG TTT GAT AAA AGT AAA ACT GTA TAT GAA GGG CCA TTT GCT TCC Tyr Leu Phe Asp Lys Ser Lys Thr Val Tyr Glu Gly Pro Phe Ala Ser 385 390 395 400	1200
AGA AGT TTA TCT GAT TGT GTA AAT TAT ATT GTA CAG GAC AGC AAA ATA Arg Ser Leu Ser Asp Cys Val Asn Tyr Ile Val Gln Asp Ser Lys Ile 405 410 415	1248
CAG CTT CCA ATT ATA CAG CTG CGT AAA GTG TGG GCT GAA GCA GTG CAC Gln Leu Pro Ile Ile Gln Leu Arg Lys Val Trp Ala Glu Ala Val His 420 425 430	1296
TAT GTG TCT GGA CTA AAA GAA GAC TAT AGC AGG CTC TTT CAG GGA CAA Tyr Val Ser Gly Leu Lys Glu Asp Tyr Ser Arg Leu Phe Gln Gly Gln 435 440 445	1344
AGG GCA GCA ATG TTA AGT CTT CTT AGA TAT AAT GCT AAC TTA ACA AAA Arg Ala Ala Met Leu Ser Leu Leu Arg Tyr Asn Ala Asn Leu Thr Lys 450 455 460	1392
ATG AAG AAC ACT TTG ATC TCA GCA TCA CAA CAA CTG AAA GCT AAA TTG Met Lys Asn Thr Leu Ile Ser Ala Ser Gln Gln Leu Lys Ala Lys Leu 465 470 475 480	1440
GAG TTT TTT CAC AAA AGC ATT CAG CTT GAC TTG GAG AGA TAC AGC GAG Glu Phe Phe His Lys Ser Ile Gln Leu Asp Leu Glu Arg Tyr Ser Glu 485 490 495	1488
CAG ATG ACG TAT GGG ATA TCT TCA GAA AAA ATG CTA AAA GCA TGG AAA Gln Met Thr Tyr Gly Ile Ser Ser Glu Lys Met Leu Lys Ala Trp Lys 500 505 510	1536
GAA ATG GAA GAA AAG GCC ATC CAC TAT GCT GAG GTT GGT GTC ATT GGA	1584

Glu Met Glu Glu Lys Ala Ile His Tyr Ala Glu Val Gly Val Ile Gly	
515 520 525	
TAC CTG GAG GAT CAG ATT ATG TCT TTG CAT GCT GAA ATC ATG GGG CTA	1632
Tyr Leu Glu Asp Gln Ile Met Ser Leu His Ala Glu Ile Met Gly Leu	
530 535 540	
CAG AAG AGC CCC TAT GGA AGA CGT CAG GGA GAC TTG ATG GAA TCT CTG	1680
Gln Lys Ser Pro Tyr Gly Arg Arg Gln Gly Asp Leu Met Glu Ser Leu	
545 550 555 560	
GAA CAG CGT GCC ATT GAT CTA TAT AAG CAG TTA AAA CAC AGA CCT TCA	1728
Glu Gln Arg Ala Ile Asp Leu Tyr Lys Gln Leu Lys His Arg Pro Ser	
565 570 575	
GAT CAC TCC TAC AGT GAC AGC ACA GAG ATG GTG AAA ATC ATT GTG CAC	1776
Asp His Ser Tyr Ser Asp Ser Thr Glu Met Val Lys Ile Ile Val His	
580 585 590	
ACT GTG CAG AGT CAG GAC CGT GTG CTC AAG GAG CTG TTT GGT CAT TTG	1824
Thr Val Gln Ser Gln Asp Arg Val Leu Lys Glu Leu Phe Gly His Leu	
595 600 605	
AGC AAG TTG TTG GGC TGT AAG CAG AAG ATT ATT GAT CTA CTC CCT AAG	1872
Ser Lys Leu Leu Gly Cys Lys Gln Lys Ile Ile Asp Leu Leu Pro Lys	
610 615 620	
GTG GAA GTG GCC CTC AGT AAT ATC AAA GAA GCT GAC AAT ACT GTC ATG	1920
Val Glu Val Ala Leu Ser Asn Ile Lys Glu Ala Asp Asn Thr Val Met	
625 630 635 640	
TTC ATG CAG GGA AAA AGG CAG AAA GAA ATA TGG CAT CTC CTT AAA ATT	1968
Phe Met Gln Gly Lys Arg Gln Lys Glu Ile Trp His Leu Leu Lys Ile	
645 650 655	
GCC TGT ACA CAG AGT TCT GCC CGC TCT CTT GTA GGA TCC AGT CTA GAA	2016
Ala Cys Thr Gln Ser Ser Ala Arg Ser Leu Val Gly Ser Ser Leu Glu	
660 665 670	
GGT GCA GTA ACC CCT CAG ACA TCA GCA TGG CTG CCC CCG ACT TCA GCA	2064
Gly Ala Val Thr Pro Gln Thr Ser Ala Trp Leu Pro Pro Thr Ser Ala	
675 680 685	
GAA CAT GAT CAT TCT CTG TCA TGT GTG GTA ACT CCT CAA GAT GGG GAG	2112
Glu His Asp His Ser Leu Ser Cys Val Val Thr Pro Gln Asp Gly Glu	
690 695 700	
ACT TCA GCA CAA ATG ATA GAA GAA AAT TTG AAC TGC CTT GGC CAT TTA	2160
Thr Ser Ala Gln Met Ile Glu Glu Asn Leu Asn Cys Leu Gly His Leu	
705 710 715 720	
AGC ACT ATT ATT CAT GAG GCA AAT GAG GAA CAG GGC AAT AGT ATG ATG	2208
Ser Thr Ile Ile His Glu Ala Asn Glu Glu Gln Gly Asn Ser Met Met	
725 730 735	
AAT CTT GAT TGG AGT TGG TTA ACA GAA TGG GTA CCG CGG GCC CGG GAT	2256
Asn Leu Asp Trp Ser Trp Leu Thr Glu Trp Val Pro Arg Ala Arg Asp	
740 745 750	

CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly 755 760 765	2304
GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys 770 775 780	2352
TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu 785 790 795 800	2400
ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro 805 810 815	2448
ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr 820 825 830	2496
CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu 835 840 845	2544
GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr 850 855 860	2592
AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg 865 870 875 880	2640
ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly 885 890 895	2688
CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala 900 905 910	2736
GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn 915 920 925	2784
ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr 930 935 940	2832
CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser 945 950 955 960	2880
ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met 965 970 975	2928
GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC	2976

Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp
 980 985 990

GAG CTG TAC AAG TAA
 Glu Leu Tyr Lys
 995

2991

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 996 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

Met	Glu	Arg	Pro	Pro	Gly	Leu	Arg	Pro	Gly	Ala	Gly	Gly	Pro	Trp	Glu	1	5	10	15
Met	Arg	Glu	Arg	Leu	Gly	Thr	Gly	Gly	Phe	Gly	Asn	Val	Cys	Leu	Tyr	20	25	30	
Gln	His	Arg	Glu	Leu	Asp	Leu	Lys	Ile	Ala	Ile	Lys	Ser	Cys	Arg	Leu	35	40	45	
Glu	Leu	Ser	Thr	Lys	Asn	Arg	Glu	Arg	Trp	Cys	His	Glu	Ile	Gln	Ile	50	55	60	
Met	Lys	Lys	Leu	Asn	His	Ala	Asn	Val	Val	Lys	Ala	Cys	Asp	Val	Pro	65	70	75	80
Glu	Glu	Leu	Asn	Ile	Leu	Ile	His	Asp	Val	Pro	Leu	Leu	Ala	Met	Glu	85	90	95	
Tyr	Cys	Ser	Gly	Gly	Asp	Leu	Arg	Lys	Leu	Leu	Asn	Lys	Pro	Glu	Asn	100	105	110	
Cys	Cys	Gly	Leu	Lys	Glu	Ser	Gln	Ile	Leu	Ser	Leu	Leu	Ser	Asp	Ile	115	120	125	
Gly	Ser	Gly	Ile	Arg	Tyr	Leu	His	Glu	Asn	Lys	Ile	Ile	His	Arg	Asp	130	135	140	
Leu	Lys	Pro	Glu	Asn	Ile	Val	Leu	Gln	Asp	Val	Gly	Gly	Lys	Ile	Ile	145	150	155	160
His	Lys	Ile	Ile	Asp	Leu	Gly	Tyr	Ala	Lys	Asp	Val	Asp	Gln	Gly	Ser	165	170	175	
Leu	Cys	Thr	Ser	Phe	Val	Gly	Thr	Leu	Gln	Tyr	Leu	Ala	Pro	Glu	Leu	180	185	190	
Phe	Glu	Asn	Lys	Pro	Tyr	Thr	Ala	Thr	Val	Asp	Tyr	Trp	Ser	Phe	Gly	195	200	205	
Thr	Met	Val	Phe	Glu	Cys	Ile	Ala	Gly	Tyr	Arg	Pro	Phe	Leu	His	His	210	215	220	
Leu	Gln	Pro	Phe	Thr	Trp	His	Glu	Lys	Ile	Lys	Lys	Lys	Asp	Pro	Lys	225	230	235	240
Cys	Ile	Phe	Ala	Cys	Glu	Glu	Met	Ser	Gly	Glu	Val	Arg	Phe	Ser	Ser	245	250	255	
His	Leu	Pro	Gln	Pro	Asn	Ser	Leu	Cys	Ser	Leu	Ile	Val	Glu	Pro	Met	260	265	270	
Glu	Asn	Trp	Leu	Gln	Leu	Met	Leu	Asn	Trp	Asp	Pro	Gln	Gln	Arg	Gly	275	280	285	
Gly	Pro	Val	Asp	Leu	Thr	Leu	Lys	Gln	Pro	Arg	Cys	Phe	Val	Leu	Met				

290	295	300
Asp His Ile Leu Asn Leu Lys Ile Val His Ile Leu Asn Met Thr Ser		
305	310	315
Ala Lys Ile Ile Ser Phe Leu Leu Pro Pro Asp Glu Ser Leu His Ser		320
	325	330
Leu Gln Ser Arg Ile Glu Arg Glu Thr Gly Ile Asn Thr Gly Ser Gln		335
	340	345
Glu Leu Leu Ser Glu Thr Gly Ile Ser Leu Asp Pro Arg Lys Pro Ala		350
	355	360
Ser Gln Cys Val Leu Asp Gly Val Arg Gly Cys Asp Ser Tyr Met Val		365
	370	375
Tyr Leu Phe Asp Lys Ser Lys Thr Val Tyr Glu Gly Pro Phe Ala Ser		380
385	390	395
Arg Ser Leu Ser Asp Cys Val Asn Tyr Ile Val Gln Asp Ser Lys Ile		400
	405	410
Gln Leu Pro Ile Ile Gln Leu Arg Lys Val Trp Ala Glu Ala Val His		415
	420	425
Tyr Val Ser Gly Leu Lys Glu Asp Tyr Ser Arg Leu Phe Gln Gly Gln		430
	435	440
Arg Ala Ala Met Leu Ser Leu Leu Arg Tyr Asn Ala Asn Leu Thr Lys		445
	450	455
Met Lys Asn Thr Leu Ile Ser Ala Ser Gln Gln Leu Lys Ala Lys Leu		460
465	470	475
Glu Phe Phe His Lys Ser Ile Gln Leu Asp Leu Glu Arg Tyr Ser Glu		480
	485	490
Gln Met Thr Tyr Gly Ile Ser Ser Glu Lys Met Leu Lys Ala Trp Lys		495
	500	505
Glu Met Glu Glu Lys Ala Ile His Tyr Ala Glu Val Gly Val Ile Gly		510
	515	520
Tyr Leu Glu Asp Gln Ile Met Ser Leu His Ala Glu Ile Met Gly Leu		525
	530	535
Gln Lys Ser Pro Tyr Gly Arg Arg Gln Gly Asp Leu Met Glu Ser Leu		540
545	550	555
Glu Gln Arg Ala Ile Asp Leu Tyr Lys Gln Leu Lys His Arg Pro Ser		560
	565	570
Asp His Ser Tyr Ser Asp Ser Thr Glu Met Val Lys Ile Ile Val His		575
	580	585
Thr Val Gln Ser Gln Asp Arg Val Leu Lys Glu Leu Phe Gly His Leu		590
	595	600
Ser Lys Leu Leu Gly Cys Lys Gln Lys Ile Ile Asp Leu Leu Pro Lys		605
	610	615
Val Glu Val Ala Leu Ser Asn Ile Lys Glu Ala Asp Asn Thr Val Met		620
625	630	635
Phe Met Gln Gly Lys Arg Gln Lys Glu Ile Trp His Leu Leu Lys Ile		640
	645	650
Ala Cys Thr Gln Ser Ser Ala Arg Ser Leu Val Gly Ser Ser Leu Glu		655
	660	665
Gly Ala Val Thr Pro Gln Thr Ser Ala Trp Leu Pro Pro Thr Ser Ala		670
	675	680
Glu His Asp His Ser Leu Ser Cys Val Val Thr Pro Gln Asp Gly Glu		685
	690	695
Thr Ser Ala Gln Met Ile Glu Glu Asn Leu Asn Cys Leu Gly His Leu		700
705	710	715
Ser Thr Ile Ile His Glu Ala Asn Glu Glu Gln Gly Asn Ser Met Met		720
	725	730
Asn Leu Asp Trp Ser Trp Leu Thr Glu Trp Val Pro Arg Ala Arg Asp		735
	740	745
Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly		750


```

      755              760              765
Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys
  770              775              780
Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu
  785              790              795              800
Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro
      805              810              815
Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr
      820              825              830
Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu
      835              840              845
Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr
      850              855              860
Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg
  865              870              875              880
Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly
      885              890              895
His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala
      900              905              910
Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn
      915              920              925
Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr
  930              935              940
Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser
  945              950              955              960
Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met
      965              970              975
Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp
      980              985              990
Glu Leu Tyr Lys
      995

```

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1908 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1905
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

```

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG      48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
  1              5              10              15

GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC      96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
      20              25              30

GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC      144

```

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG	336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
100 105 110	
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC	384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly	
115 120 125	
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC	432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr	
130 135 140	
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC	480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn	
145 150 155 160	
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC	528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser	
165 170 175	
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC	576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly	
180 185 190	
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG	624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu	
195 200 205	
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC	672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe	
210 215 220	
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC	720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser	
225 230 235 240	
GGA CTC AGA TCT CGA GCT CAA GCT TCC ATG AGC GAG ACG GTC ATC ATG	768
Gly Leu Arg Ser Arg Ala Gln Ala Ser Met Ser Glu Thr Val Ile Met	
245 250 255	
AGC GAG ACG GTC ATC TGT TCC AGC CGG GCC ACT GTG ATG CTT TAT GAT	816
Ser Glu Thr Val Ile Cys Ser Ser Arg Ala Thr Val Met Leu Tyr Asp	
260 265 270	

GAT GGC AAC AAG CGA TGG CTC CCT GCT GGC ACG GGT CCC CAG GCC TTC Asp Gly Asn Lys Arg Trp Leu Pro Ala Gly Thr Gly Pro Gln Ala Phe 275 280 285	864
AGC CGC GTC CAG ATC TAC CAC AAC CCC ACG GCC AAT TCC TTT CGC GTC Ser Arg Val Gln Ile Tyr His Asn Pro Thr Ala Asn Ser Phe Arg Val 290 295 300	912
GTG GGC CGG AAG ATG CAG CCC GAC CAG CAG GTG GTC ATC AAC TGT GCC Val Gly Arg Lys Met Gln Pro Asp Gln Gln Val Val Ile Asn Cys Ala 305 310 315 320	960
ATC GTC CGG GGT GTC AAG TAT AAC CAG GCC ACC CCC AAC TTC CAT CAG Ile Val Arg Gly Val Lys Tyr Asn Gln Ala Thr Pro Asn Phe His Gln 325 330 335	1008
TGG CGC GAC GCT CGC CAG GTC TGG GGC CTC AAC TTC GGC AGC AAG GAG Trp Arg Asp Ala Arg Gln Val Trp Gly Leu Asn Phe Gly Ser Lys Glu 340 345 350	1056
GAT GCG GCC CAG TTT GCC GCC GGC ATG GCC AGT GCC CTA GAG GCG TTG Asp Ala Ala Gln Phe Ala Ala Gly Met Ala Ser Ala Leu Glu Ala Leu 355 360 365	1104
GAA GGA GGT GGG CCC CCT CCA CCC CCA GCA CTT CCC ACC TGG TCG GTC Glu Gly Gly Gly Pro Pro Pro Pro Pro Ala Leu Pro Thr Trp Ser Val 370 375 380	1152
CCG AAC GGC CCC TCC CCG GAG GAG GTG GAG CAG CAG AAA AGG CAG CAG Pro Asn Gly Pro Ser Pro Glu Glu Val Glu Gln Gln Lys Arg Gln Gln 385 390 395 400	1200
CCC GGC CCG TCG GAG CAC ATA GAG CGC CGG GTC TCC AAT GCA GGA GGC Pro Gly Pro Ser Glu His Ile Glu Arg Arg Val Ser Asn Ala Gly Gly 405 410 415	1248
CCA CCT GCT CCC CCC GCT GGG GGT CCA CCC CCA CCA CCA GGA CCT CCC Pro Pro Ala Pro Pro Ala Gly Gly Pro Pro Pro Pro Pro Gly Pro Pro 420 425 430	1296
CCT CCT CCA GGT CCC CCC CCA CCC CCA GGT TTG CCC CCT TCG GGG GTC Pro Pro Pro Gly Pro Pro Pro Pro Pro Gly Leu Pro Pro Ser Gly Val 435 440 445	1344
CCA GCT GCA GCG CAC GGA GCA GGG GGA GGA CCA CCC CCT GCA CCC CCT Pro Ala Ala Ala His Gly Ala Gly Gly Gly Pro Pro Pro Ala Pro Pro 450 455 460	1392
CTC CCG GCA GCA CAG GGC CCT GGT GGT GGG GGA GCT GGG GCC CCA GGC Leu Pro Ala Ala Gln Gly Pro Gly Gly Gly Gly Ala Gly Ala Pro Gly 465 470 475 480	1440
CTG GCC GCA GCT ATT GCT GGA GCC AAA CTC AGG AAA GTC AGC AAG CAG Leu Ala Ala Ala Ile Ala Gly Ala Lys Leu Arg Lys Val Ser Lys Gln 485 490 495	1488
GAG GAG GCC TCA GGG GGG CCC ACA GCC CCC AAA GCT GAG AGT GGT CGA	1536

Glu	Glu	Ala	Ser	Gly	Gly	Pro	Thr	Ala	Pro	Lys	Ala	Glu	Ser	Gly	Arg		
			500					505					510				
AGC	GGA	GGT	GGG	GGA	CTC	ATG	GAA	GAG	ATG	AAC	GCC	ATG	CTG	GCC	CGG	1584	
Ser	Gly	Gly	Gly	Gly	Leu	Met	Glu	Glu	Met	Asn	Ala	Met	Leu	Ala	Arg		
		515					520				525						
AGA	AGG	AAA	GCC	ACG	CAA	GTT	GGG	GAG	AAA	ACC	CCC	AAG	GAT	GAA	TCT	1632	
Arg	Arg	Lys	Ala	Thr	Gln	Val	Gly	Glu	Lys	Thr	Pro	Lys	Asp	Glu	Ser		
		530				535					540						
GCC	AAT	CAG	GAG	GAG	CCA	GAG	GCC	AGA	GTC	CCG	GCC	CAG	AGT	GAA	TCT	1680	
Ala	Asn	Gln	Glu	Glu	Pro	Glu	Ala	Arg	Val	Pro	Ala	Gln	Ser	Glu	Ser		
545					550					555				560			
GTG	CGG	AGA	CCC	TGG	GAG	AAG	AAC	AGC	ACA	ACC	TTG	CCA	AGG	ATG	AAG	1728	
Val	Arg	Arg	Pro	Trp	Glu	Lys	Asn	Ser	Thr	Thr	Leu	Pro	Arg	Met	Lys		
				565					570					575			
TCG	TCT	TCT	TCG	GTG	ACC	ACT	TCC	GAG	ACC	CAA	CCC	TGC	ACG	CCC	AGC	1776	
Ser	Ser	Ser	Ser	Val	Thr	Thr	Ser	Glu	Thr	Gln	Pro	Cys	Thr	Pro	Ser		
				580				585					590				
TCC	AGT	GAT	TAC	TCG	GAC	CTA	CAG	AGG	GTG	AAA	CAG	GAG	CTT	CTG	GAA	1824	
Ser	Ser	Asp	Tyr	Ser	Asp	Leu	Gln	Arg	Val	Lys	Gln	Glu	Leu	Leu	Glu		
		595				600					605						
GAG	GTG	AAG	AAG	GAA	TTG	CAG	AAA	GTG	AAA	GAG	GAA	ATC	ATT	GAA	GCC	1872	
Glu	Val	Lys	Lys	Glu	Leu	Gln	Lys	Val	Lys	Glu	Glu	Ile	Ile	Glu	Ala		
		610				615					620						
TTC	GTC	CAG	GAG	CTG	AGG	AAG	CGG	GGT	TCT	CCC	TGA					1908	
Phe	Val	Gln	Glu	Leu	Arg	Lys	Arg	Gly	Ser	Pro							
625					630					635							

(2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 635 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu		
1				5					10					15			
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly		
			20					25					30				
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile		
		35				40						45					
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr		
		50				55					60						
Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys		

65		70		75		80
Gln His Asp Phe	Phe Lys Ser Ala Met Pro	Glu Gly Tyr Val	Gln Glu			
	85	90	95			
Arg Thr Ile Phe	Phe Lys Asp Asp Gly Asn Tyr Lys Thr	Arg Ala Glu				
	100	105	110			
Val Lys Phe Glu	Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly					
	115	120	125			
Ile Asp Phe Lys	Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr					
	130	135	140			
Asn Tyr Asn Ser His	Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn					
	145	150	155			160
Gly Ile Lys Val	Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser					
	165	170	175			
Val Gln Leu Ala	Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly					
	180	185	190			
Pro Val Leu Leu	Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu					
	195	200	205			
Ser Lys Asp Pro	Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe					
	210	215	220			
Val Thr Ala Ala	Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser					
	225	230	235			240
Gly Leu Arg Ser	Arg Ala Gln Ala Ser Met Ser Glu Thr Val Ile Met					
	245	250	255			
Ser Glu Thr Val	Ile Cys Ser Ser Arg Ala Thr Val Met Leu Tyr Asp					
	260	265	270			
Asp Gly Asn Lys	Arg Trp Leu Pro Ala Gly Thr Gly Pro Gln Ala Phe					
	275	280	285			
Ser Arg Val Gln	Ile Tyr His Asn Pro Thr Ala Asn Ser Phe Arg Val					
	290	295	300			
Val Gly Arg Lys	Met Gln Pro Asp Gln Gln Val Val Ile Asn Cys Ala					
	305	310	315			320
Ile Val Arg Gly	Val Lys Tyr Asn Gln Ala Thr Pro Asn Phe His Gln					
	325	330	335			
Trp Arg Asp Ala	Arg Gln Val Trp Gly Leu Asn Phe Gly Ser Lys Glu					
	340	345	350			
Asp Ala Ala Gln	Phe Ala Ala Gly Met Ala Ser Ala Leu Glu Ala Leu					
	355	360	365			
Glu Gly Gly Gly	Pro Pro Pro Pro Ala Leu Pro Thr Trp Ser Val					
	370	375	380			
Pro Asn Gly Pro	Ser Pro Glu Val Glu Gln Gln Lys Arg Gln Gln					
	385	390	395			400
Pro Gly Pro Ser	Glu His Ile Glu Arg Arg Val Ser Asn Ala Gly Gly					
	405	410	415			
Pro Pro Ala Pro	Pro Ala Gly Gly Pro Pro Pro Pro Pro Gly Pro Pro					
	420	425	430			
Pro Pro Pro Gly	Pro Pro Pro Pro Gly Leu Pro Pro Ser Gly Val					
	435	440	445			
Pro Ala Ala Ala	His Gly Ala Gly Gly Gly Pro Pro Pro Ala Pro Pro					
	450	455	460			
Leu Pro Ala Ala	Gln Gly Pro Gly Gly Gly Gly Ala Gly Ala Pro Gly					
	465	470	475			480
Leu Ala Ala Ala	Ile Ala Gly Ala Lys Leu Arg Lys Val Ser Lys Gln					
	485	490	495			
Glu Glu Ala Ser	Gly Gly Pro Thr Ala Pro Lys Ala Glu Ser Gly Arg					
	500	505	510			
Ser Gly Gly Gly	Gly Gly Leu Met Glu Glu Met Asn Ala Met Leu Ala Arg					
	515	520	525			
Arg Arg Lys Ala	Thr Gln Val Gly Glu Lys Thr Pro Lys Asp Glu Ser					

```

      530              535              540
Ala Asn Gln Glu Glu Pro Glu Ala Arg Val Pro Ala Gln Ser Glu Ser
545              550              555              560
Val Arg Arg Pro Trp Glu Lys Asn Ser Thr Thr Leu Pro Arg Met Lys
      565              570              575
Ser Ser Ser Ser Val Thr Thr Ser Glu Thr Gln Pro Cys Thr Pro Ser
      580              585              590
Ser Ser Asp Tyr Ser Asp Leu Gln Arg Val Lys Gln Glu Leu Leu Glu
      595              600              605
Glu Val Lys Lys Glu Leu Gln Lys Val Lys Glu Glu Ile Ile Glu Ala
      610              615              620
Phe Val Gln Glu Leu Arg Lys Arg Gly Ser Pro
625              630              635

```

(2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1329 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1326
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

```

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG      48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
  1              5              10              15

GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC      96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
      20              25              30

GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC      144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
      35              40              45

TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC      192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
      50              55              60

CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG      240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
      65              70              75              80

CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG      288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
      85              90              95

CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG      336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
      100              105              110

```

GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly 115 120 125	384
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr 130 135 140	432
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn 145 150 155 160	480
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser 165 170 175	528
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190	576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205	624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220	672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser 225 230 235 240	720
GGA CTC AGA TCT CGA GCT CAA GCT TCA ATG GCT GCC ATC CGG AAG AAA Gly Leu Arg Ser Arg Ala Gln Ala Ser Met Ala Ala Ile Arg Lys Lys 245 250 255	768
CTG GTG ATT GTT GGT GAT GGA GCC TGT GGA AAG ACA TGC TTG CTC ATA Leu Val Ile Val Gly Asp Gly Ala Cys Gly Lys Thr Cys Leu Leu Ile 260 265 270	816
GTC TTC AGC AAG GAC CAG TTC CCA GAG GTG TAT GTG CCC ACA GTG TTT Val Phe Ser Lys Asp Gln Phe Pro Glu Val Tyr Val Pro Thr Val Phe 275 280 285	864
GAG AAC TAT GTG GCA GAT ATC GAG GTG GAT GGA AAG CAG GTA GAG TTG Glu Asn Tyr Val Ala Asp Ile Glu Val Asp Gly Lys Gln Val Glu Leu 290 295 300	912
GCT TTG TGG GAC ACA GCT GGG CAG GAA GAT TAT GAT CGC CTG AGG CCC Ala Leu Trp Asp Thr Ala Gly Gln Glu Asp Tyr Asp Arg Leu Arg Pro 305 310 315 320	960
CTC TCC TAC CCA GAT ACC GAT GTT ATA CTG ATG TGT TTT TCC ATC GAC Leu Ser Tyr Pro Asp Thr Asp Val Ile Leu Met Cys Phe Ser Ile Asp 325 330 335	1008
AGC CCT GAT AGT TTA GAA AAC ATC CCA GAA AAG TGG ACC CCA GAA GTC	1056

Ser Pro Asp Ser Leu Glu Asn Ile Pro Glu Lys Trp Thr Pro Glu Val
 340 345 350

AAG CAT TTC TGT CCC AAC GTG CCC ATC ATC CTG GTT GGG AAT AAG AAG 1104
 Lys His Phe Cys Pro Asn Val Pro Ile Ile Leu Val Gly Asn Lys Lys
 355 360 365

GAT CTT CGG AAT GAT GAG CAC ACA AGG CGG GAG CTA GCC AAG ATG AAG 1152
 Asp Leu Arg Asn Asp Glu His Thr Arg Arg Glu Leu Ala Lys Met Lys
 370 375 380

CAG GAG CCG GTG AAA CCT GAA GAA GGC AGA GAT ATG GCA AAC AGG ATT 1200
 Gln Glu Pro Val Lys Pro Glu Glu Gly Arg Asp Met Ala Asn Arg Ile
 385 390 395 400

GGC GCT TTT GGG TAC ATG GAG TGT TCA GCA AAG ACC AAA GAT GGA GTG 1248
 Gly Ala Phe Gly Tyr Met Glu Cys Ser Ala Lys Thr Lys Asp Gly Val
 405 410 415

AGA GAG GTT TTT GAA ATG GCT ACG AGA GCT GCT CTG CAA GCT AGA CGT 1296
 Arg Glu Val Phe Glu Met Ala Thr Arg Ala Ala Leu Gln Ala Arg Arg
 420 425 430

GGG AAG AAA AAA TCT GGT TGC CTT GTC TTG TGA 1329
 Gly Lys Lys Lys Ser Gly Cys Leu Val Leu
 435 440

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 442 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10 15
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 20 25 30
 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 35 40 45
 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50 55 60
 Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 65 70 75 80
 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85 90 95
 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 100 105 110
 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 115 120 125
 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr

130		135		140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn				
145		150		155
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser				160
	165		170	175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly				
	180		185	190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu				
	195		200	205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe				
	210		215	220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser				
225		230		235
Gly Leu Arg Ser Arg Ala Gln Ala Ser Met Ala Ala Ile Arg Lys Lys				240
	245		250	255
Leu Val Ile Val Gly Asp Gly Ala Cys Gly Lys Thr Cys Leu Leu Ile				
	260		265	270
Val Phe Ser Lys Asp Gln Phe Pro Glu Val Tyr Val Pro Thr Val Phe				
	275		280	285
Glu Asn Tyr Val Ala Asp Ile Glu Val Asp Gly Lys Gln Val Glu Leu				
	290		295	300
Ala Leu Trp Asp Thr Ala Gly Gln Glu Asp Tyr Asp Arg Leu Arg Pro				
305		310		315
Leu Ser Tyr Pro Asp Thr Asp Val Ile Leu Met Cys Phe Ser Ile Asp				
	325		330	335
Ser Pro Asp Ser Leu Glu Asn Ile Pro Glu Lys Trp Thr Pro Glu Val				
	340		345	350
Lys His Phe Cys Pro Asn Val Pro Ile Ile Leu Val Gly Asn Lys Lys				
	355		360	365
Asp Leu Arg Asn Asp Glu His Thr Arg Arg Glu Leu Ala Lys Met Lys				
	370		375	380
Gln Glu Pro Val Lys Pro Glu Glu Gly Arg Asp Met Ala Asn Arg Ile				
385		390		395
Gly Ala Phe Gly Tyr Met Glu Cys Ser Ala Lys Thr Lys Asp Gly Val				
	405		410	415
Arg Glu Val Phe Glu Met Ala Thr Arg Ala Ala Leu Gln Ala Arg Arg				
	420		425	430
Gly Lys Lys Lys Ser Gly Cys Leu Val Leu				
	435		440	

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1140 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1137
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

ATG GAC CAT TAT GAT TCT CAG CAA ACC AAC GAT TAC ATG CAG CCA GAA

Met	Asp	His	Tyr	Asp	Ser	Gln	Gln	Thr	Asn	Asp	Tyr	Met	Gln	Pro	Glu	
1				5					10					15		
GAG	GAC	TGG	GAC	CGG	GAC	CTG	CTC	CTG	GAC	CCG	GCC	TGG	GAG	AAG	CAG	96
Glu	Asp	Trp	Asp	Arg	Asp	Leu	Leu	Leu	Asp	Pro	Ala	Trp	Glu	Lys	Gln	
			20					25					30			
CAG	AGA	AAG	ACA	TTC	ACG	GCA	TGG	TGT	AAC	TCC	CAC	CTC	CGG	AAG	GCG	144
Gln	Arg	Lys	Thr	Phe	Thr	Ala	Trp	Cys	Asn	Ser	His	Leu	Arg	Lys	Ala	
		35					40					45				
GGG	ACA	CAG	ATC	GAG	AAC	ATC	GAA	GAG	GAC	TTC	CGG	GAT	GGC	CTG	AAG	192
Gly	Thr	Gln	Ile	Glu	Asn	Ile	Glu	Glu	Asp	Phe	Arg	Asp	Gly	Leu	Lys	
	50					55				60						
CTC	ATG	CTG	CTG	CTG	GAG	GTC	ATC	TCA	GGT	GAA	CGC	TTG	GCC	AAG	CCA	240
Leu	Met	Leu	Leu	Leu	Glu	Val	Ile	Ser	Gly	Glu	Arg	Leu	Ala	Lys	Pro	
65					70					75					80	
GAG	CGA	GGC	AAG	ATG	AGA	GTG	CAC	AAG	ATC	TCC	AAC	GTC	AAC	AAG	GCC	288
Glu	Arg	Gly	Lys	Met	Arg	Val	His	Lys	Ile	Ser	Asn	Val	Asn	Lys	Ala	
				85					90					95		
CTG	GAT	TTC	ATA	GCC	AGC	AAA	GGC	GTC	AAA	CTG	GTG	TCC	ATC	GGA	GCC	336
Leu	Asp	Phe	Ile	Ala	Ser	Lys	Gly	Val	Lys	Leu	Val	Ser	Ile	Gly	Ala	
		100						105					110			
GAA	GAA	ATC	CTG	GAT	GGG	AAT	GTG	AAG	ATG	ACC	CTG	GGC	ATG	ATC	TGG	384
Glu	Glu	Ile	Val	Asp	Gly	Asn	Val	Lys	Met	Thr	Leu	Gly	Met	Ile	Trp	
		115					120					125				
ACC	ATC	ATC	CTG	CGC	AGG	GAT	CCA	CCG	GTC	GCC	ACC	ATG	GTG	AGC	AAG	432
Thr	Ile	Ile	Leu	Arg	Arg	Asp	Pro	Pro	Val	Ala	Thr	Met	Val	Ser	Lys	
		130				135						140				
GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	GTC	GAG	CTG	GAC	480
Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	Val	Glu	Leu	Asp	
145					150					155					160	
GGC	GAC	GTA	AAC	GAC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	GAG	GGC	GAG	GGC	528
Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	Glu	Gly	Glu	Gly	
			165						170					175		
GAT	GCC	ACC	TAC	GAC	AAG	CTG	ACC	CTG	AAG	TTC	ATC	TGC	ACC	ACC	GGC	576
Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	Thr	Thr	Gly	
			180					185					190			
AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	CTG	ACC	TAC	GGC	624
Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	Leu	Thr	Tyr	Gly	
		195					200						205			
GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	CAC	ATG	AAG	CAG	CAC	GAC	TTC	672
Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	Gln	His	Asp	Phe	
	210					215					220					
TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	CGC	ACC	ATC	TTC	720
Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	Arg	Thr	Ile	Phe	
225					230					235					240	

TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu 245 250 255	768
GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys 260 265 270	816
GAG GAC GGC AAC ATC CTG GGC CAC AAG CTG GAG TAC AAC TAC AAC AGC Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser 275 280 285	864
CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val 290 295 300	912
AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala 305 310 315 320	960
GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu 325 330 335	1008
CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro 340 345 350	1056
AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala 355 360 365	1104
GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys 370 375	1140

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 379 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

Met Asp His Tyr Asp Ser Gln Gln Thr Asn Asp Tyr Met Gln Pro Glu 1 5 10 15
Glu Asp Trp Asp Arg Asp Leu Leu Leu Asp Pro Ala Trp Glu Lys Gln 20 25 30
Gln Arg Lys Thr Phe Thr Ala Trp Cys Asn Ser His Leu Arg Lys Ala 35 40 45
Gly Thr Gln Ile Glu Asn Ile Glu Glu Asp Phe Arg Asp Gly Leu Lys

50		55		60	
Leu Met Leu Leu Leu Glu Val Ile Ser Gly Glu Arg Leu Ala Lys Pro					
65		70		75	80
Glu Arg Gly Lys Met Arg Val His Lys Ile Ser Asn Val Asn Lys Ala					
	85		90		95
Leu Asp Phe Ile Ala Ser Lys Gly Val Lys Leu Val Ser Ile Gly Ala					
	100		105		110
Glu Glu Ile Val Asp Gly Asn Val Lys Met Thr Leu Gly Met Ile Trp					
	115		120		125
Thr Ile Ile Leu Arg Arg Asp Pro Pro Val Ala Thr Met Val Ser Lys					
	130		135		140
Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp					
145		150		155	160
Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly					
	165		170		175
Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly					
	180		185		190
Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly					
	195		200		205
Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe					
	210		215		220
Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe					
225		230		235	240
Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu					
	245		250		255
Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys					
	260		265		270
Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser					
	275		280		285
His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val					
	290		295		300
Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala					
305		310		315	320
Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu					
	325		330		335
Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro					
	340		345		350
Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala					
	355		360		365
Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys					
370		375			

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3516 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...3513
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr	
50 55 60	
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG	240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys	
65 70 75 80	
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG	288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu	
85 90 95	
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG	336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu	
100 105 110	
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC	384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly	
115 120 125	
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGC CAC AAG CTG GAG TAC	432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr	
130 135 140	
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC	480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn	
145 150 155 160	
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC	528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser	
165 170 175	
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC	576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly	
180 185 190	
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG	624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu	
195 200 205	
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC	672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe	
210 215 220	
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC	720

Val Gln Leu His Gly Tyr Met Glu Asn Lys Pro Leu Gly Leu Gln Ile	
690 695 700	
TTC ATT GGG ACA GCT GAT GAG CGG ATC CTT AAG CCG CAC GCC TTC TAC	2160
Phe Ile Gly Thr Ala Asp Glu Arg Ile Leu Lys Pro His Ala Phe Tyr	
705 710 715 720	
CAG GTG CAC CGA ATC ACG GGG AAA ACT GTC ACC ACC ACC AGC TAT GAG	2208
Gln Val His Arg Ile Thr Gly Lys Thr Val Thr Thr Thr Ser Tyr Glu	
725 730 735	
AAG ATA GTG GGC AAC ACC AAA GTC CTG GAG ATC CCC TTG GAG CCC AAA	2256
Lys Ile Val Gly Asn Thr Lys Val Leu Glu Ile Pro Leu Glu Pro Lys	
740 745 750	
AAC AAC ATG AGG GCA ACC ATC GAC TGT GCG GGG ATC TTG AAG CTT AGA	2304
Asn Asn Met Arg Ala Thr Ile Asp Cys Ala Gly Ile Leu Lys Leu Arg	
755 760 765	
AAC GCC GAC ATT GAG CTG CGG AAA GGC GAG ACG GAC ATT GGA AGA AAG	2352
Asn Ala Asp Ile Glu Leu Arg Lys Gly Glu Thr Asp Ile Gly Arg Lys	
770 775 780	
AAC ACG CGG GTG AGA CTG GTT TTC CGA GTT CAC ATC CCA GAG TCC AGT	2400
Asn Thr Arg Val Arg Leu Val Phe Arg Val His Ile Pro Glu Ser Ser	
785 790 795 800	
GGC AGA ATC GTC TCT TTA CAG ACT GCA TCT AAC CCC ATC GAG TGC TCC	2448
Gly Arg Ile Val Ser Leu Gln Thr Ala Ser Asn Pro Ile Glu Cys Ser	
805 810 815	
CAG CGA TCT GCT CAC GAG CTG CCC ATG GTT GAA AGA CAA GAC ACA GAC	2496
Gln Arg Ser Ala His Glu Leu Pro Met Val Glu Arg Gln Asp Thr Asp	
820 825 830	
AGC TGC CTG GTC TAT GGC GGC CAG CAA ATG ATC CTC ACG GGG CAG AAC	2544
Ser Cys Leu Val Tyr Gly Gly Gln Gln Met Ile Leu Thr Gly Gln Asn	
835 840 845	
TTT ACA TCC GAG TCC AAA GTT GTG TTT ACT GAG AAG ACC ACA GAT GGA	2592
Phe Thr Ser Glu Ser Lys Val Val Phe Thr Glu Lys Thr Thr Asp Gly	
850 855 860	
CAG CAA ATT TGG GAG ATG GAA GCC ACG GTG GAT AAG GAC AAG AGC CAG	2640
Gln Gln Ile Trp Glu Met Glu Ala Thr Val Asp Lys Asp Lys Ser Gln	
865 870 875 880	
CCC AAC ATG CTT TTT GTT GAG ATC CCT GAA TAT CGG AAC AAG CAT ATC	2688
Pro Asn Met Leu Phe Val Glu Ile Pro Glu Tyr Arg Asn Lys His Ile	
885 890 895	
CGC ACA CCT GTA AAA GTG AAC TTC TAC GTC ATC AAT GGG AAG AGA AAA	2736
Arg Thr Pro Val Lys Val Asn Phe Tyr Val Ile Asn Gly Lys Arg Lys	
900 905 910	
CGA AGT CAG CCT CAG CAC TTT ACC TAC CAC CCA GTC CCA GCC ATC AAG	2784
Arg Ser Gln Pro Gln His Phe Thr Tyr His Pro Val Pro Ala Ile Lys	
915 920 925	

ACG GAG CCC ACG GAT GAA TAT GAC CCC ACT CTG ATC TGC AGC CCC ACC Thr Glu Pro Thr Asp Glu Tyr Asp Pro Thr Leu Ile Cys Ser Pro Thr 930 935 940	2832
CAT GGA GGC CTG GGG AGC CAG CCT TAC TAC CCC CAG CAC CCG ATG GTG His Gly Gly Leu Gly Ser Gln Pro Tyr Tyr Pro Gln His Pro Met Val 945 950 955 960	2880
GCC GAG TCC CCC TCC TGC CTC GTG GCC ACC ATG GCT CCC TGC CAG CAG Ala Glu Ser Pro Ser Cys Leu Val Ala Thr Met Ala Pro Cys Gln Gln 965 970 975	2928
TTC CGC ACG GGG CTC TCA TCC CCT GAC GCC CGC TAC CAG CAA CAG AAC Phe Arg Thr Gly Leu Ser Ser Pro Asp Ala Arg Tyr Gln Gln Gln Asn 980 985 990	2976
CCA GCG GCC GTA CTC TAC CAG CGG AGC AAG AGC CTG AGC CCC AGC CTG Pro Ala Ala Val Leu Tyr Gln Arg Ser Lys Ser Leu Ser Pro Ser Leu 995 1000 1005	3024
CTG GGC TAT CAG CAG CCG GCC CTC ATG GCC GCC CCG CTG TCC CTT GCG Leu Gly Tyr Gln Gln Pro Ala Leu Met Ala Ala Pro Leu Ser Leu Ala 1010 1015 1020	3072
GAC GCT CAC CGC TCT GTG CTG GTG CAC GCC GGC TCC CAG GGC CAG AGC Asp Ala His Arg Ser Val Leu Val His Ala Gly Ser Gln Gly Gln Ser 1025 1030 1035 1040	3120
TCA GCC CTG CTC CAC CCC TCT CCG ACC AAC CAG CAG GCC TCG CCT GTG Ser Ala Leu Leu His Pro Ser Pro Thr Asn Gln Gln Ala Ser Pro Val 1045 1050 1055	3168
ATC CAC TAC TCA CCC ACC AAC CAG CAG CTG CGC TGC GGA AGC CAC CAG Ile His Tyr Ser Pro Thr Asn Gln Gln Leu Arg Cys Gly Ser His Gln 1060 1065 1070	3216
GAG TTC CAG CAC ATC ATG TAC TGC GAG AAT TTC GCA CCA GGC ACC ACC Glu Phe Gln His Ile Met Tyr Cys Glu Asn Phe Ala Pro Gly Thr Thr 1075 1080 1085	3264
AGA CCT GGC CCG CCC CCG GTC AGT CAA GGT CAG AGG CTG AGC CCG GGT Arg Pro Gly Pro Pro Pro Val Ser Gln Gly Gln Arg Leu Ser Pro Gly 1090 1095 1100	3312
TCC TAC CCC ACA GTC ATT CAG CAG CAG AAT GCC ACG AGC CAA AGA GCC Ser Tyr Pro Thr Val Ile Gln Gln Gln Asn Ala Thr Ser Gln Arg Ala 1105 1110 1115 1120	3360
GCC AAA AAC GGA CCC CCG GTC AGT GAC CAA AAG GAA GTA TTA CCT GCG Ala Lys Asn Gly Pro Pro Val Ser Asp Gln Lys Glu Val Leu Pro Ala 1125 1130 1135	3408
GGG GTG ACC ATT AAA CAG GAG CAG AAC TTG GAC CAG ACC TAC TTG GAT Gly Val Thr Ile Lys Gln Glu Gln Asn Leu Asp Gln Thr Tyr Leu Asp 1140 1145 1150	3456
GAT GTT AAT GAA ATT ATC AGG AAG GAG TTT TCA GGA CCT CCT GCC AGA	3504

Asp Val Asn Glu Ile Ile Arg Lys Glu Phe Ser Gly Pro Pro Ala Arg
 1155 1160 1165

AAT CAG ACG TAA
 Asn Gln Thr
 1170

3516

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1171 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

```

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1           5           10           15
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
      20           25           30
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
      35           40           45
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
      50           55           60
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
      65           70           75           80
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
      85           90           95
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
      100          105          110
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
      115          120          125
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
      130          135          140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
      145          150          155          160
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
      165          170          175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
      180          185          190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
      195          200          205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
      210          215          220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
      225          230          235          240
Gly Leu Arg Ser Arg Ala Met Asn Ala Pro Glu Arg Gln Pro Gln Pro
      245          250          255
Asp Gly Gly Asp Ala Pro Gly His Glu Pro Gly Gly Ser Pro Gln Asp
      260          265          270
Glu Leu Asp Phe Ser Ile Leu Phe Asp Tyr Glu Tyr Leu Asn Pro Asn
      275          280          285
Glu Glu Glu Pro Asn Ala His Lys Val Ala Ser Pro Pro Ser Gly Pro

```

290	295	300
Ala Tyr Pro Asp Asp Val Met Asp Tyr Gly Leu Lys Pro Tyr Ser Pro		
305	310	315
Leu Ala Ser Leu Ser Gly Glu Pro Pro Gly Arg Phe Gly Glu Pro Asp		320
	325	330
Arg Val Gly Pro Gln Lys Phe Leu Ser Ala Ala Lys Pro Ala Gly Ala		335
	340	345
Ser Gly Leu Ser Pro Arg Ile Glu Ile Thr Pro Ser His Glu Leu Ile		350
	355	360
Gln Ala Val Gly Pro Leu Arg Met Arg Asp Ala Gly Leu Leu Val Glu		365
	370	375
Gln Pro Pro Leu Ala Gly Val Ala Ala Ser Pro Arg Phe Thr Leu Pro		380
385	390	395
Val Pro Gly Phe Glu Gly Tyr Arg Glu Pro Leu Cys Leu Ser Pro Ala		400
	405	410
Ser Ser Gly Ser Ser Ala Ser Phe Ile Ser Asp Thr Phe Ser Pro Tyr		415
	420	425
Thr Ser Pro Cys Val Ser Pro Asn Asn Gly Gly Pro Asp Leu Cys		430
	435	440
Pro Gln Phe Gln Asn Ile Pro Ala His Tyr Ser Pro Arg Thr Ser Pro		445
	450	455
Ile Met Ser Pro Arg Thr Ser Leu Ala Glu Asp Ser Cys Leu Gly Arg		460
465	470	475
His Ser Pro Val Pro Arg Pro Ala Ser Arg Ser Ser Ser Pro Gly Ala		480
	485	490
Lys Arg Arg His Ser Cys Ala Glu Ala Leu Val Ala Leu Pro Pro Gly		495
	500	505
Ala Ser Pro Gln Arg Ser Arg Ser Pro Ser Pro Gln Pro Ser Ser His		510
	515	520
Val Ala Pro Gln Asp His Gly Ser Pro Ala Gly Tyr Pro Pro Val Ala		525
	530	535
Gly Ser Ala Val Ile Met Asp Ala Leu Asn Ser Leu Ala Thr Asp Ser		540
545	550	555
Pro Cys Gly Ile Pro Pro Lys Met Trp Lys Thr Ser Pro Asp Pro Ser		560
	565	570
Pro Val Ser Ala Ala Pro Ser Lys Ala Gly Leu Pro Arg His Ile Tyr		575
	580	585
Pro Ala Val Glu Phe Leu Gly Pro Cys Glu Gln Gly Glu Arg Arg Asn		590
	595	600
Ser Ala Pro Glu Ser Ile Leu Leu Val Pro Pro Thr Trp Pro Lys Pro		605
	610	615
Leu Val Pro Ala Ile Pro Ile Cys Ser Ile Pro Val Thr Ala Ser Leu		620
625	630	635
Pro Pro Leu Glu Trp Pro Leu Ser Ser Gln Ser Gly Ser Tyr Glu Leu		640
	645	650
Arg Ile Glu Val Gln Pro Lys Pro His His Arg Ala His Tyr Glu Thr		655
	660	665
Glu Gly Ser Arg Gly Ala Val Lys Ala Pro Thr Gly Gly His Pro Val		670
	675	680
Val Gln Leu His Gly Tyr Met Glu Asn Lys Pro Leu Gly Leu Gln Ile		685
	690	695
Phe Ile Gly Thr Ala Asp Glu Arg Ile Leu Lys Pro His Ala Phe Tyr		700
705	710	715
Gln Val His Arg Ile Thr Gly Lys Thr Val Thr Thr Thr Ser Tyr Glu		720
	725	730
Lys Ile Val Gly Asn Thr Lys Val Leu Glu Ile Pro Leu Glu Pro Lys		735
	740	745
Asn Asn Met Arg Ala Thr Ile Asp Cys Ala Gly Ile Leu Lys Leu Arg		750

755		760		765
Asn Ala Asp Ile Glu Leu Arg Lys Gly Glu Thr Asp Ile Gly Arg Lys				
770		775		780
Asn Thr Arg Val Arg Leu Val Phe Arg Val His Ile Pro Glu Ser Ser				
785		790		800
Gly Arg Ile Val Ser Leu Gln Thr Ala Ser Asn Pro Ile Glu Cys Ser				
	805		810	815
Gln Arg Ser Ala His Glu Leu Pro Met Val Glu Arg Gln Asp Thr Asp				
	820		825	830
Ser Cys Leu Val Tyr Gly Gly Gln Gln Met Ile Leu Thr Gly Gln Asn				
	835		840	845
Phe Thr Ser Glu Ser Lys Val Val Phe Thr Glu Lys Thr Thr Asp Gly				
	850		855	860
Gln Gln Ile Trp Glu Met Glu Ala Thr Val Asp Lys Asp Lys Ser Gln				
865		870		880
Pro Asn Met Leu Phe Val Glu Ile Pro Glu Tyr Arg Asn Lys His Ile				
	885		890	895
Arg Thr Pro Val Lys Val Asn Phe Tyr Val Ile Asn Gly Lys Arg Lys				
	900		905	910
Arg Ser Gln Pro Gln His Phe Thr Tyr His Pro Val Pro Ala Ile Lys				
	915		920	925
Thr Glu Pro Thr Asp Glu Tyr Asp Pro Thr Leu Ile Cys Ser Pro Thr				
	930		935	940
His Gly Gly Leu Gly Ser Gln Pro Tyr Tyr Pro Gln His Pro Met Val				
945		950		960
Ala Glu Ser Pro Ser Cys Leu Val Ala Thr Met Ala Pro Cys Gln Gln				
	965		970	975
Phe Arg Thr Gly Leu Ser Ser Pro Asp Ala Arg Tyr Gln Gln Gln Asn				
	980		985	990
Pro Ala Ala Val Leu Tyr Gln Arg Ser Lys Ser Leu Ser Pro Ser Leu				
	995		1000	1005
Leu Gly Tyr Gln Gln Pro Ala Leu Met Ala Ala Pro Leu Ser Leu Ala				
	1010		1015	1020
Asp Ala His Arg Ser Val Leu Val His Ala Gly Ser Gln Gly Gln Ser				
025		1030		1040
Ser Ala Leu Leu His Pro Ser Pro Thr Asn Gln Gln Ala Ser Pro Val				
	1045		1050	1055
Ile His Tyr Ser Pro Thr Asn Gln Gln Leu Arg Cys Gly Ser His Gln				
	1060		1065	1070
Glu Phe Gln His Ile Met Tyr Cys Glu Asn Phe Ala Pro Gly Thr Thr				
	1075		1080	1085
Arg Pro Gly Pro Pro Pro Val Ser Gln Gly Gln Arg Leu Ser Pro Gly				
	1090		1095	1100
Ser Tyr Pro Thr Val Ile Gln Gln Gln Asn Ala Thr Ser Gln Arg Ala				
105		1110		1120
Ala Lys Asn Gly Pro Pro Val Ser Asp Gln Lys Glu Val Leu Pro Ala				
	1125		1130	1135
Gly Val Thr Ile Lys Gln Glu Gln Asn Leu Asp Gln Thr Tyr Leu Asp				
	1140		1145	1150
Asp Val Asn Glu Ile Ile Arg Lys Glu Phe Ser Gly Pro Pro Ala Arg				
	1155		1160	1165
Asn Gln Thr				
1170				

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3546 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA
 (ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 1...3543
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

ATG AAC GCC CCC GAG CGG CAG CCC CAA CCC GAC GGC GGG GAC GCC CCA	48
Met Asn Ala Pro Glu Arg Gln Pro Gln Pro Asp Gly Gly Asp Ala Pro	
1 5 10 15	
GGC CAC GAG CCT GGG GGC AGC CCC CAA GAC GAG CTT GAC TTC TCC ATC	96
Gly His Glu Pro Gly Gly Ser Pro Gln Asp Glu Leu Asp Phe Ser Ile	
20 25 30	
CTC TTC GAC TAT GAG TAT TTG AAT CCG AAC GAA GAA GAG CCG AAT GCA	144
Leu Phe Asp Tyr Glu Tyr Leu Asn Pro Asn Glu Glu Glu Pro Asn Ala	
35 40 45	
CAT AAG GTC GCC AGC CCA CCC TCC GGA CCC GCA TAC CCC GAT GAT GTA	192
His Lys Val Ala Ser Pro Pro Ser Gly Pro Ala Tyr Pro Asp Asp Val	
50 55 60	
ATG GAC TAT GGC CTC AAG CCA TAC AGC CCC CTT GCT AGT CTC TCT GGC	240
Met Asp Tyr Gly Leu Lys Pro Tyr Ser Pro Leu Ala Ser Leu Ser Gly	
65 70 75 80	
GAG CCC CCC GGC CGA TTC GGA GAG CCG GAT AGG GTA GGG CCG CAG AAG	288
Glu Pro Pro Gly Arg Phe Gly Glu Pro Asp Arg Val Gly Pro Gln Lys	
85 90 95	
TTT CTG AGC GCG GCC AAG CCA GCA GGG GCC TCG GGC CTG AGC CCT CGG	336
Phe Leu Ser Ala Ala Lys Pro Ala Gly Ala Ser Gly Leu Ser Pro Arg	
100 105 110	
ATC GAG ATC ACT CCG TCC CAC GAA CTG ATC CAG GCA GTG GGG CCC CTC	384
Ile Glu Ile Thr Pro Ser His Glu Leu Ile Gln Ala Val Gly Pro Leu	
115 120 125	
CGC ATG AGA GAC GCG GGC CTC CTG GTG GAG CAG CCT CCC CTG GCC GGC	432
Arg Met Arg Asp Ala Gly Leu Leu Val Glu Gln Pro Pro Leu Ala Gly	
130 135 140	
GTG GCC GCC AGC CCG AGG TTC ACC CTG CCC GTG CCC GGC TTC GAG GGC	480
Val Ala Ala Ser Pro Arg Phe Thr Leu Pro Val Pro Gly Phe Glu Gly	
145 150 155 160	
TAC CGC GAG CCG CTT TGC TTG AGC CCC GCT AGC AGC GGC TCC TCT GCC	528
Tyr Arg Glu Pro Leu Cys Leu Ser Pro Ala Ser Ser Gly Ser Ser Ala	
165 170 175	
AGC TTC ATT TCT GAC ACC TTC TCC CCC TAC ACC TCG CCC TGC GTC TCG	576

Ser Phe Ile Ser Asp Thr Phe Ser Pro Tyr Thr Ser Pro Cys Val Ser	
180 185 190	
CCC AAT AAC GGC GGC CCC GAC GAC CTG TGT CCG CAG TTT CAA AAC ATC	624
Pro Asn Asn Gly Gly Pro Asp Asp Leu Cys Pro Gln Phe Gln Asn Ile	
195 200 205	
CCT GCT CAT TAT TCC CCC AGA ACC TCG CCA ATA ATG TCA CCT CGA ACC	672
Pro Ala His Tyr Ser Pro Arg Thr Ser Pro Ile Met Ser Pro Arg Thr	
210 215 220	
AGC CTC GCC GAG GAC AGC TGC CTG GGC CGC CAC TCG CCC GTG CCC CGT	720
Ser Leu Ala Glu Asp Ser Cys Leu Gly Arg His Ser Pro Val Pro Arg	
225 230 235 240	
CCG GCC TCC CGC TCC TCA TCG CCT GGT GCC AAG CGG AGG CAT TCG TGC	768
Pro Ala Ser Arg Ser Ser Ser Pro Gly Ala Lys Arg Arg His Ser Cys	
245 250 255	
GCC GAG GCC TTG GTT GCC CTG CCG CCC GGA GCC TCA CCC CAG CGC TCC	816
Ala Glu Ala Leu Val Ala Leu Pro Pro Gly Ala Ser Pro Gln Arg Ser	
260 265 270	
CGG AGC CCC TCG CCG CAG CCC TCA TCT CAC GTG GCA CCC CAG GAC CAC	864
Arg Ser Pro Ser Pro Gln Pro Ser Ser His Val Ala Pro Gln Asp His	
275 280 285	
GGC TCC CCG GCT GGG TAC CCC CCT GTG GCT GGC TCT GCC GTG ATC ATG	912
Gly Ser Pro Ala Gly Tyr Pro Pro Val Ala Gly Ser Ala Val Ile Met	
290 295 300	
GAT GCC CTG AAC AGC CTC GCC ACG GAC TCG CCT TGT GGG ATC CCC CCC	960
Asp Ala Leu Asn Ser Leu Ala Thr Asp Ser Pro Cys Gly Ile Pro Pro	
305 310 315 320	
AAG ATG TGG AAG ACC AGC CCT GAC CCC TCG CCG GTG TCT GCC GCC CCA	1008
Lys Met Trp Lys Thr Ser Pro Asp Pro Ser Pro Val Ser Ala Ala Pro	
325 330 335	
TCC AAG GCC GGC CTG CCT CGC CAC ATC TAC CCG GCC GTG GAG TTC CTG	1056
Ser Lys Ala Gly Leu Pro Arg His Ile Tyr Pro Ala Val Glu Phe Leu	
340 345 350	
GGG CCC TGC GAG CAG GGC GAG AGG AGA AAC TCG GCT CCA GAA TCC ATC	1104
Gly Pro Cys Glu Gln Gly Glu Arg Arg Asn Ser Ala Pro Glu Ser Ile	
355 360 365	
CTG CTG GTT CCG CCC ACT TGG CCC AAG CCG CTG GTG CCT GCC ATT CCC	1152
Leu Leu Val Pro Pro Thr Trp Pro Lys Pro Leu Val Pro Ala Ile Pro	
370 375 380	
ATC TGC AGC ATC CCA GTG ACT GCA TCC CTC CCT CCA CTT GAG TGG CCG	1200
Ile Cys Ser Ile Pro Val Thr Ala Ser Leu Pro Pro Leu Glu Trp Pro	
385 390 395 400	
CTG TCC AGT CAG TCA GGC TCT TAC GAG CTG CGG ATC GAG GTG CAG CCC	1248
Leu Ser Ser Gln Ser Gly Ser Tyr Glu Leu Arg Ile Glu Val Gln Pro	
405 410 415	

AAG CCA CAT CAC CGG GCC CAC TAT GAG ACA GAA GGC AGC CGA GGG GCT	1296
Lys Pro His His Arg Ala His Tyr Glu Thr Glu Gly Ser Arg Gly Ala	
420 425 430	
GTC AAA GCT CCA ACT GGA GGC CAC CCT GTG GTT CAG CTC CAT GGC TAC	1344
Val Lys Ala Pro Thr Gly Gly His Pro Val Val Gln Leu His Gly Tyr	
435 440 445	
ATG GAA AAC AAG CCT CTG GGA CTT CAG ATC TTC ATT GGG ACA GCT GAT	1392
Met Glu Asn Lys Pro Leu Gly Leu Gln Ile Phe Ile Gly Thr Ala Asp	
450 455 460	
GAG CGG ATC CTT AAG CCG CAC GCC TTC TAC CAG GTG CAC CGA ATC ACG	1440
Glu Arg Ile Leu Lys Pro His Ala Phe Tyr Gln Val His Arg Ile Thr	
465 470 475 480	
GGG AAA ACT GTC ACC ACC ACC AGC TAT GAG AAG ATA GTG GGC AAC ACC	1488
Gly Lys Thr Val Thr Thr Thr Ser Tyr Glu Lys Ile Val Gly Asn Thr	
485 490 495	
AAA GTC CTG GAG ATC CCC TTG GAG CCC AAA AAC AAC ATG AGG GCA ACC	1536
Lys Val Leu Glu Ile Pro Leu Glu Pro Lys Asn Asn Met Arg Ala Thr	
500 505 510	
ATC GAC TGT GCG GGG ATC TTG AAG CTT AGA AAC GCC GAC ATT GAG CTG	1584
Ile Asp Cys Ala Gly Ile Leu Lys Leu Arg Asn Ala Asp Ile Glu Leu	
515 520 525	
CGG AAA GGC GAG ACG GAC ATT GGA AGA AAG AAC ACG CGG GTG AGA CTG	1632
Arg Lys Gly Glu Thr Asp Ile Gly Arg Lys Asn Thr Arg Val Arg Leu	
530 535 540	
GTT TTC CGA GTT CAC ATC CCA GAG TCC AGT GGC AGA ATC GTC TCT TTA	1680
Val Phe Arg Val His Ile Pro Glu Ser Ser Gly Arg Ile Val Ser Leu	
545 550 555 560	
CAG ACT GCA TCT AAC CCC ATC GAG TGC TCC CAG CGA TCT GCT CAC GAG	1728
Gln Thr Ala Ser Asn Pro Ile Glu Cys Ser Gln Arg Ser Ala His Glu	
565 570 575	
CTG CCC ATG GTT GAA AGA CAA GAC ACA GAC AGC TGC CTG GTC TAT GGC	1776
Leu Pro Met Val Glu Arg Gln Asp Thr Asp Ser Cys Leu Val Tyr Gly	
580 585 590	
GGC CAG CAA ATG ATC CTC ACG GGG CAG AAC TTT ACA TCC GAG TCC AAA	1824
Gly Gln Gln Met Ile Leu Thr Gly Gln Asn Phe Thr Ser Glu Ser Lys	
595 600 605	
GTT GTG TTT ACT GAG AAG ACC ACA GAT GGA CAG CAA ATT TGG GAG ATG	1872
Val Val Phe Thr Glu Lys Thr Thr Asp Gly Gln Gln Ile Trp Glu Met	
610 615 620	
GAA GCC ACG GTG GAT AAG GAC AAG AGC CAG CCC AAC ATG CTT TTT GTT	1920
Glu Ala Thr Val Asp Lys Asp Lys Ser Gln Pro Asn Met Leu Phe Val	
625 630 635 640	
GAG ATC CCT GAA TAT CGG AAC AAG CAT ATC CGC ACA CCT GTA AAA GTG	1968

Glu Ile Pro Glu Tyr Arg Asn Lys His Ile Arg Thr Pro Val Lys Val	
645 650 655	
AAC TTC TAC GTC ATC AAT GGG AAG AGA AAA CGA AGT CAG CCT CAG CAC	2016
Asn Phe Tyr Val Ile Asn Gly Lys Arg Lys Arg Ser Gln Pro Gln His	
660 665 670	
TTT ACC TAC CAC CCA GTC CCA GCC ATC AAG ACG GAG CCC ACG GAT GAA	2064
Phe Thr Tyr His Pro Val Pro Ala Ile Lys Thr Glu Pro Thr Asp Glu	
675 680 685	
TAT GAC CCC ACT CTG ATC TGC AGC CCC ACC CAT GGA GGC CTG GGG AGC	2112
Tyr Asp Pro Thr Leu Ile Cys Ser Pro Thr His Gly Gly Leu Gly Ser	
690 695 700	
CAG CCT TAC TAC CCC CAG CAC CCG ATG GTG GCC GAG TCC CCC TCC TGC	2160
Gln Pro Tyr Tyr Pro Gln His Pro Met Val Ala Glu Ser Pro Ser Cys	
705 710 715 720	
CTC GTG GCC ACC ATG GCT CCC TGC CAG CAG TTC CGC ACG GGG CTC TCA	2208
Leu Val Ala Thr Met Ala Pro Cys Gln Gln Phe Arg Thr Gly Leu Ser	
725 730 735	
TCC CCT GAC GCC CGC TAC CAG CAA CAG AAC CCA GCG GCC GTA CTC TAC	2256
Ser Pro Asp Ala Arg Tyr Gln Gln Asn Pro Ala Ala Val Leu Tyr	
740 745 750	
CAG CGG AGC AAG AGC CTG AGC CCC AGC CTG CTG GGC TAT CAG CAG CCG	2304
Gln Arg Ser Lys Ser Leu Ser Pro Ser Leu Leu Gly Tyr Gln Gln Pro	
755 760 765	
GCC CTC ATG GCC GCC CCG CTG TCC CTT GCG GAC GCT CAC CGC TCT GTG	2352
Ala Leu Met Ala Ala Pro Leu Ser Leu Ala Asp Ala His Arg Ser Val	
770 775 780	
CTG GTG CAC GCC GGC TCC CAG GGC CAG AGC TCA GCC CTG CTC CAC CCC	2400
Leu Val His Ala Gly Ser Gln Gly Gln Ser Ser Ala Leu Leu His Pro	
785 790 795 800	
TCT CCG ACC AAC CAG CAG GCC TCG CCT GTG ATC CAC TAC TCA CCC ACC	2448
Ser Pro Thr Asn Gln Gln Ala Ser Pro Val Ile His Tyr Ser Pro Thr	
805 810 815	
AAC CAG CAG CTG CGC TGC GGA AGC CAC CAG GAG TTC CAG CAC ATC ATG	2496
Asn Gln Gln Leu Arg Cys Gly Ser His Gln Glu Phe Gln His Ile Met	
820 825 830	
TAC TGC GAG AAT TTC GCA CCA GGC ACC ACC AGA CCT GGC CCG CCC CCG	2544
Tyr Cys Glu Asn Phe Ala Pro Gly Thr Thr Arg Pro Gly Pro Pro Pro	
835 840 845	
GTC AGT CAA GGT CAG AGG CTG AGT CCG GGT TCC TAC CCC ACA GTC ATT	2592
Val Ser Gln Gly Gln Arg Leu Ser Pro Gly Ser Tyr Pro Thr Val Ile	
850 855 860	
CAG CAG CAG AAT GCC ACG AGC CAA AGA GCC GCC AAA AAC GGA CCC CCG	2640
Gln Gln Gln Asn Ala Thr Ser Gln Arg Ala Lys Asn Gly Pro Pro	
865 870 875 880	

GTC AGT GAC CAA AAG GAA GTA TTA CCT GCG GGG GTG ACC ATT AAA CAG	2688
Val Ser Asp Gln Lys Glu Val Leu Pro Ala Gly Val Thr Ile Lys Gln	
885 890 895	
GAG CAG AAC TTG GAC CAG ACC TAC TTG GAT GAT GTT AAT GAA ATT ATC	2736
Glu Gln Asn Leu Asp Gln Thr Tyr Leu Asp Asp Val Asn Glu Ile Ile	
900 905 910	
AGG AAG GAG TTT TCA GGA CCT CCT GCC AGA AAT CAG ACG AGA ATT CTG	2784
Arg Lys Glu Phe Ser Gly Pro Pro Ala Arg Asn Gln Thr Arg Ile Leu	
915 920 925	
CAG TCG ACG GTA CCG CGG GCC CGG GAT CCA CCG GTC GCC ACC ATG GTG	2832
Gln Ser Thr Val Pro Arg Ala Arg Asp Pro Pro Val Ala Thr Met Val	
930 935 940	
AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC GAG	2880
Ser Lys Gly Glu Glu Phe Thr Gly Val Val Pro Ile Leu Val Glu	
945 950 955 960	
CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC	2928
Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly	
965 970 975	
GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC	2976
Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr	
980 985 990	
ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG ACC	3024
Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr	
995 1000 1005	
TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC	3072
Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His	
1010 1015 1020	
GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC	3120
Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr	
1025 1030 1035 1040	
ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG	3168
Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys	
1045 1050 1055	
TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC	3216
Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp	
1060 1065 1070	
TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC	3264
Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr	
1075 1080 1085	
AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC	3312
Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile	
1090 1095 1100	
AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG	3360

Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln
 1105 1110 1115 1120
 CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG 3408
 Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val
 1125 1130 1135
 CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA 3456
 Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys
 1140 1145 1150
 GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC 3504
 Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr
 1155 1160 1165
 GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA 3546
 Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
 1170 1175 1180

(2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1181 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

Met Asn Ala Pro Glu Arg Gln Pro Gln Pro Asp Gly Gly Asp Ala Pro
 1 5 10 15
 Gly His Glu Pro Gly Gly Ser Pro Gln Asp Glu Leu Asp Phe Ser Ile
 20 25 30
 Leu Phe Asp Tyr Glu Tyr Leu Asn Pro Asn Glu Glu Glu Pro Asn Ala
 35 40 45
 His Lys Val Ala Ser Pro Pro Ser Gly Pro Ala Tyr Pro Asp Asp Val
 50 55 60
 Met Asp Tyr Gly Leu Lys Pro Tyr Ser Pro Leu Ala Ser Leu Ser Gly
 65 70 75 80
 Glu Pro Pro Gly Arg Phe Gly Glu Pro Asp Arg Val Gly Pro Gln Lys
 85 90 95
 Phe Leu Ser Ala Ala Lys Pro Ala Gly Ala Ser Gly Leu Ser Pro Arg
 100 105 110
 Ile Glu Ile Thr Pro Ser His Glu Leu Ile Gln Ala Val Gly Pro Leu
 115 120 125
 Arg Met Arg Asp Ala Gly Leu Val Glu Gln Pro Pro Leu Ala Gly
 130 135 140
 Val Ala Ala Ser Pro Arg Phe Thr Leu Pro Val Pro Gly Phe Glu Gly
 145 150 155 160
 Tyr Arg Glu Pro Leu Cys Leu Ser Pro Ala Ser Ser Gly Ser Ser Ala
 165 170 175
 Ser Phe Ile Ser Asp Thr Phe Ser Pro Tyr Thr Ser Pro Cys Val Ser
 180 185 190
 Pro Asn Asn Gly Gly Pro Asp Asp Leu Cys Pro Gln Phe Gln Asn Ile

195	200	205
Pro Ala His Tyr Ser Pro Arg Thr Ser Pro Ile Met Ser Pro Arg Thr		
210	215	220
Ser Leu Ala Glu Asp Ser Cys Leu Gly Arg His Ser Pro Val Pro Arg		
225	230	235
Pro Ala Ser Arg Ser Ser Ser Pro Gly Ala Lys Arg Arg His Ser Cys		
245	250	255
Ala Glu Ala Leu Val Ala Leu Pro Pro Gly Ala Ser Pro Gln Arg Ser		
260	265	270
Arg Ser Pro Ser Pro Gln Pro Ser Ser His Val Ala Pro Gln Asp His		
275	280	285
Gly Ser Pro Ala Gly Tyr Pro Pro Val Ala Gly Ser Ala Val Ile Met		
290	295	300
Asp Ala Leu Asn Ser Leu Ala Thr Asp Ser Pro Cys Gly Ile Pro Pro		
305	310	315
Lys Met Trp Lys Thr Ser Pro Asp Pro Ser Pro Val Ser Ala Ala Pro		
325	330	335
Ser Lys Ala Gly Leu Pro Arg His Ile Tyr Pro Ala Val Glu Phe Leu		
340	345	350
Gly Pro Cys Glu Gln Gly Glu Arg Arg Asn Ser Ala Pro Glu Ser Ile		
355	360	365
Leu Leu Val Pro Pro Thr Trp Pro Lys Pro Leu Val Pro Ala Ile Pro		
370	375	380
Ile Cys Ser Ile Pro Val Thr Ala Ser Leu Pro Pro Leu Glu Trp Pro		
385	390	395
Leu Ser Ser Gln Ser Gly Ser Tyr Glu Leu Arg Ile Glu Val Gln Pro		
405	410	415
Lys Pro His His Arg Ala His Tyr Glu Thr Glu Gly Ser Arg Gly Ala		
420	425	430
Val Lys Ala Pro Thr Gly Gly His Pro Val Val Gln Leu His Gly Tyr		
435	440	445
Met Glu Asn Lys Pro Leu Gly Leu Gln Ile Phe Ile Gly Thr Ala Asp		
450	455	460
Glu Arg Ile Leu Lys Pro His Ala Phe Tyr Gln Val His Arg Ile Thr		
465	470	475
Gly Lys Thr Val Thr Thr Thr Ser Tyr Glu Lys Ile Val Gly Asn Thr		
485	490	495
Lys Val Leu Glu Ile Pro Leu Glu Pro Lys Asn Asn Met Arg Ala Thr		
500	505	510
Ile Asp Cys Ala Gly Ile Leu Lys Leu Arg Asn Ala Asp Ile Glu Leu		
515	520	525
Arg Lys Gly Glu Thr Asp Ile Gly Arg Lys Asn Thr Arg Val Arg Leu		
530	535	540
Val Phe Arg Val His Ile Pro Glu Ser Ser Gly Arg Ile Val Ser Leu		
545	550	555
Gln Thr Ala Ser Asn Pro Ile Glu Cys Ser Gln Arg Ser Ala His Glu		
565	570	575
Leu Pro Met Val Glu Arg Gln Asp Thr Asp Ser Cys Leu Val Tyr Gly		
580	585	590
Gly Gln Gln Met Ile Leu Thr Gly Gln Asn Phe Thr Ser Glu Ser Lys		
595	600	605
Val Val Phe Thr Glu Lys Thr Thr Asp Gly Gln Gln Ile Trp Glu Met		
610	615	620
Glu Ala Thr Val Asp Lys Asp Lys Ser Gln Pro Asn Met Leu Phe Val		
625	630	635
Glu Ile Pro Glu Tyr Arg Asn Lys His Ile Arg Thr Pro Val Lys Val		
645	650	655
Asn Phe Tyr Val Ile Asn Gly Lys Arg Lys Arg Ser Gln Pro Gln His		

660	665	670
Phe Thr Tyr His Pro Val Pro Ala Ile Lys Thr Glu Pro Thr Asp Glu		
675	680	685
Tyr Asp Pro Thr Leu Ile Cys Ser Pro Thr His Gly Gly Leu Gly Ser		
690	695	700
Gln Pro Tyr Tyr Pro Gln His Pro Met Val Ala Glu Ser Pro Ser Cys		
705	710	715
Leu Val Ala Thr Met Ala Pro Cys Gln Gln Phe Arg Thr Gly Leu Ser		
725	730	735
Ser Pro Asp Ala Arg Tyr Gln Gln Gln Asn Pro Ala Ala Val Leu Tyr		
740	745	750
Gln Arg Ser Lys Ser Leu Ser Pro Ser Leu Leu Gly Tyr Gln Gln Pro		
755	760	765
Ala Leu Met Ala Ala Pro Leu Ser Leu Ala Asp Ala His Arg Ser Val		
770	775	780
Leu Val His Ala Gly Ser Gln Gly Gln Ser Ser Ala Leu Leu His Pro		
785	790	795
Ser Pro Thr Asn Gln Gln Ala Ser Pro Val Ile His Tyr Ser Pro Thr		
805	810	815
Asn Gln Gln Leu Arg Cys Gly Ser His Gln Glu Phe Gln His Ile Met		
820	825	830
Tyr Cys Glu Asn Phe Ala Pro Gly Thr Thr Arg Pro Gly Pro Pro Pro		
835	840	845
Val Ser Gln Gly Gln Arg Leu Ser Pro Gly Ser Tyr Pro Thr Val Ile		
850	855	860
Gln Gln Gln Asn Ala Thr Ser Gln Arg Ala Ala Lys Asn Gly Pro Pro		
865	870	875
Val Ser Asp Gln Lys Glu Val Leu Pro Ala Gly Val Thr Ile Lys Gln		
885	890	895
Glu Gln Asn Leu Asp Gln Thr Tyr Leu Asp Asp Val Asn Glu Ile Ile		
900	905	910
Arg Lys Glu Phe Ser Gly Pro Pro Ala Arg Asn Gln Thr Arg Ile Leu		
915	920	925
Gln Ser Thr Val Pro Arg Ala Arg Asp Pro Pro Val Ala Thr Met Val		
930	935	940
Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu		
945	950	955
Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly		
965	970	975
Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr		
980	985	990
Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr		
995	1000	1005
Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His		
1010	1015	1020
Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr		
1025	1030	1035
Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys		
1045	1050	1055
Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp		
1060	1065	1070
Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr		
1075	1080	1085
Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile		
1090	1095	1100
Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln		
1105	1110	1115
Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val		

```

          1125              1130              1135
Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys
          1140              1145              1150
Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr
          1155              1160              1165
Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
          1170              1175              1180

```

(2) INFORMATION FOR SEQ ID NO:134:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2802 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2799
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

```

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG      48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
  1              5              10              15

GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC      96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
          20              25              30

GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC      144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
          35              40              45

TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC      192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
          50              55              60

CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG      240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
          65              70              75              80

CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG      288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
          85              90              95

CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG      336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
          100             105             110

GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC      384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
          115             120             125

ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC      432

```

Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	
130						135						140				
AAC	TAC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	480
Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	
145					150					155					160	
GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	528
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	
				165					170					175		
GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	576
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	
			180					185					190			
CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	TCC	GCC	CTG	624
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	
		195					200					205				
AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	CTG	GAG	TTC	672
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	
	210					215					220					
GTG	ACC	GCC	GCC	GGG	ATC	ACT	CTC	GGC	ATG	GAC	GAG	CTG	TAC	AAG	TCC	720
Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Ser	
225				230						235					240	
GGA	CTC	AGA	TCT	CGA	GGG	AGC	ATG	GGC	ACC	TTG	CGG	GAT	TTA	CAG	TAC	768
Gly	Leu	Arg	Ser	Arg	Gly	Ser	Met	Gly	Thr	Leu	Arg	Asp	Leu	Gln	Tyr	
				245					250					255		
GCG	CTC	CAG	GAG	AAG	ATC	GAG	GAG	CTG	AGG	CAG	CGG	GAT	GCT	CTC	ATC	816
Ala	Leu	Gln	Glu	Lys	Ile	Glu	Glu	Leu	Arg	Gln	Arg	Asp	Ala	Leu	Ile	
		260						265					270			
GAC	GAG	CTG	GAG	CTG	GAG	TTG	GAT	CAG	AAG	GAC	GAA	CTG	ATC	CAG	AAG	864
Asp	Glu	Leu	Glu	Leu	Glu	Leu	Asp	Gln	Lys	Asp	Glu	Leu	Ile	Gln	Lys	
		275					280					285				
CTG	CAG	AAC	GAG	CTG	GAC	AAG	TAC	CGC	TCG	GTG	ATC	CGA	CCA	GCC	ACC	912
Leu	Gln	Asn	Glu	Leu	Asp	Lys	Tyr	Arg	Ser	Val	Ile	Arg	Pro	Ala	Thr	
		290				295					300					
CAG	CAG	GCG	CAG	AAG	CAG	AGC	GCG	AGC	ACC	TTG	CAG	GGC	GAG	CCG	CGC	960
Gln	Gln	Ala	Gln	Lys	Gln	Ser	Ala	Ser	Thr	Leu	Gln	Gly	Glu	Pro	Arg	
305					310				315						320	
ACC	AAG	CGG	CAG	GCG	ATC	TCC	GCC	GAG	CCC	ACC	GCC	TTC	GAC	ATC	CAG	1008
Thr	Lys	Arg	Gln	Ala	Ile	Ser	Ala	Glu	Pro	Thr	Ala	Phe	Asp	Ile	Gln	
				325					330					335		
GAT	CTC	AGC	CAT	GTG	ACC	CTG	CCC	TTC	TAC	CCC	AAG	AGC	CCA	CAG	TCC	1056
Asp	Leu	Ser	His	Val	Thr	Leu	Pro	Phe	Tyr	Pro	Lys	Ser	Pro	Gln	Ser	
			340					345					350			
AAG	GAT	CTT	ATA	AAG	GAA	GCT	ATC	CTT	GAC	AAT	GAC	TTT	ATG	AAG	AAC	1104
Lys	Asp	Leu	Ile	Lys	Glu	Ala	Ile	Leu	Asp	Asn	Asp	Phe	Met	Lys	Asn	
		355					360					365				

TTG GAG CTG TCG CAG ATC CAG GAG ATT GTG GAT TGT ATG TAC CCG GTG	1152
Leu Glu Leu Ser Gln Ile Gln Glu Ile Val Asp Cys Met Tyr Pro Val	
370 375 380	
GAG TAT GGC AAG GAC AGT TGC ATC ATC AAA GAA GGA GAC GTG GGG TCA	1200
Glu Tyr Gly Lys Asp Ser Cys Ile Ile Lys Glu Gly Asp Val Gly Ser	
385 390 395 400	
CTG GTG TAT GTC ATG GAA GAT GGT AAG GTT GAA GTT ACA AAA GAA GGT	1248
Leu Val Tyr Val Met Glu Asp Gly Lys Val Glu Val Thr Lys Glu Gly	
405 410 415	
CTG AAG TTG TGT ACC ATG GGT CCA GGA AAA GTG TTT GGG GAA TTG GCT	1296
Val Lys Leu Cys Thr Met Gly Pro Gly Lys Val Phe Gly Glu Leu Ala	
420 425 430	
ATT CTT TAC AAC TGT ACC CGG ACA GCG ACC GTC AAG ACT CTT GTA AAT	1344
Ile Leu Tyr Asn Cys Thr Arg Thr Ala Thr Val Lys Thr Leu Val Asn	
435 440 445	
GTA AAA CTC TGG GCC ATT GAT CGA CAA TGT TTT CAA ACA ATA ATG ATG	1392
Val Lys Leu Trp Ala Ile Asp Arg Gln Cys Phe Gln Thr Ile Met Met	
450 455 460	
AGG ACA GGA CTC ATC AAG CAT ACC GAG TAT ATG GAA TTT TTA AAA AGC	1440
Arg Thr Gly Leu Ile Lys His Thr Glu Tyr Met Glu Phe Leu Lys Ser	
465 470 475 480	
GTT CCA ACA TTC CAG AGC CTT CCT GAA GAG ATC CTC AGC AAG CTT GCT	1488
Val Pro Thr Phe Gln Ser Leu Pro Glu Glu Ile Leu Ser Lys Leu Ala	
485 490 495	
GAT GTC CTT GAA GAG ACC CAC TAT GAA AAT GGA GAA TAT ATT ATC AGG	1536
Asp Val Leu Glu Glu Thr His Tyr Glu Asn Gly Glu Tyr Ile Ile Arg	
500 505 510	
CAA GGT GCA AGA GGG GAC ACC TTC TTT ATC ATC AGC AAA GGA ACG GTA	1584
Gln Gly Ala Arg Gly Asp Thr Phe Phe Ile Ile Ser Lys Gly Thr Val	
515 520 525	
AAT GTC ACT CGT GAA GAC TCA CCG AGT GAA GAC CCA GTC TTT CTT AGA	1632
Asn Val Thr Arg Glu Asp Ser Pro Ser Glu Asp Pro Val Phe Leu Arg	
530 535 540	
ACT TTA GGA AAA GGA GAC TGG TTT GGA GAG AAA GCC TTG CAG GGG GAA	1680
Thr Leu Gly Lys Gly Asp Trp Phe Gly Glu Lys Ala Leu Gln Gly Glu	
545 550 555 560	
GAT GTG AGA ACA GCA AAC GTA ATT GCT GCA GAA GCT GTA ACC TGC CTT	1728
Asp Val Arg Thr Ala Asn Val Ile Ala Ala Glu Ala Val Thr Cys Leu	
565 570 575	
GTG ATT GAC AGA GAC TCT TTT AAA CAT TTG ATT GGA GGG CTG GAT GAT	1776
Val Ile Asp Arg Asp Ser Phe Lys His Leu Ile Gly Gly Leu Asp Asp	
580 585 590	
GTT TCT AAT AAA GCA TAT GAA GAT GCA GAA GCT AAA GCA AAA TAT GAA	1824

Val Ser Asn Lys Ala Tyr Glu Asp Ala Glu Ala Lys Ala Lys Tyr Glu	
595 600 605	
GCT GAA GCG GCT TTC TTC GCC AAC CTG AAG CTG TCT GAT TTC AAC ATC	1872
Ala Glu Ala Ala Phe Phe Ala Asn Leu Lys Leu Ser Asp Phe Asn Ile	
610 615 620	
ATT GAT ACC CTT GGA GTT GGA GGT TTC GGA CGA GTA GAA CTG GTC CAG	1920
Ile Asp Thr Leu Gly Val Gly Gly Phe Gly Arg Val Glu Leu Val Gln	
625 630 635 640	
TTG AAA AGT GAA GAA TCC AAA ACG TTT GCA ATG AAG ATT CTC AAG AAA	1968
Leu Lys Ser Glu Glu Ser Lys Thr Phe Ala Met Lys Ile Leu Lys Lys	
645 650 655	
CGT CAC ATT GTG GAC ACA AGA CAG CAG GAG CAC ATC CGC TCA GAG AAG	2016
Arg His Ile Val Asp Thr Arg Gln Gln Glu His Ile Arg Ser Glu Lys	
660 665 670	
CAG ATC ATG CAG GGG GCT CAT TCC GAT TTC ATA GTG AGA CTG TAC AGA	2064
Gln Ile Met Gln Gly Ala His Ser Asp Phe Ile Val Arg Leu Tyr Arg	
675 680 685	
ACA TTT AAG GAC AGC AAA TAT TTG TAT ATG TTG ATG GAA GCT TGT CTA	2112
Thr Phe Lys Asp Ser Lys Tyr Leu Tyr Met Leu Met Glu Ala Cys Leu	
690 695 700	
GGT GGA GAG CTC TGG ACC ATT CTC AGG GAT AGA GGT TCG TTT GAA GAT	2160
Gly Gly Glu Leu Trp Thr Ile Leu Arg Asp Arg Gly Ser Phe Glu Asp	
705 710 715 720	
TCT ACA ACC AGA TTT TAC ACA GCA TGT GTG GTA GAA GCT TTT GCC TAT	2208
Ser Thr Thr Arg Phe Tyr Thr Ala Cys Val Val Glu Ala Phe Ala Tyr	
725 730 735	
CTG CAT TCC AAA GGA ATC ATT TAC AGG GAC CTC AAG CCA GAA AAT CTC	2256
Leu His Ser Lys Gly Ile Ile Tyr Arg Asp Leu Lys Pro Glu Asn Leu	
740 745 750	
ATC CTA GAT CAC CGA GGT TAT GCC AAA CTG GTT GAT TTT GGC TTT GCA	2304
Ile Leu Asp His Arg Gly Tyr Ala Lys Leu Val Asp Phe Gly Phe Ala	
755 760 765	
AAG AAA ATA GGA TTT GGA AAG AAA ACA TGG ACT TTT TGT GGG ACT CCA	2352
Lys Lys Ile Gly Phe Gly Lys Lys Thr Trp Thr Phe Cys Gly Thr Pro	
770 775 780	
GAG TAT GTA GCC CCA GAG ATC ATC CTG AAC AAA GGC CAT GAC ATT TCA	2400
Glu Tyr Val Ala Pro Glu Ile Ile Leu Asn Lys Gly His Asp Ile Ser	
785 790 795 800	
GCC GAC TAC TGG TCA CTG GGA ATC CTA ATG TAT GAA CTC CTG ACT GGC	2448
Ala Asp Tyr Trp Ser Leu Gly Ile Leu Met Tyr Glu Leu Leu Thr Gly	
805 810 815	
AGC CCA CCT TTC TCA GGC CCA GAT CCT ATG AAA ACC TAT AAC ATC ATA	2496
Ser Pro Pro Phe Ser Gly Pro Asp Pro Met Lys Thr Tyr Asn Ile Ile	
820 825 830	

TTG AGG GGG ATT GAC ATG ATA GAA TTT CCA AAG AAG ATT GCC AAA AAT Leu Arg Gly Ile Asp Met Ile Glu Phe Pro Lys Lys Ile Ala Lys Asn 835 840 845	2544
GCT GCT AAT TTA ATT AAA AAA CTA TGC AGG GAC AAT CCA TCA GAA AGA Ala Ala Asn Leu Ile Lys Lys Leu Cys Arg Asp Asn Pro Ser Glu Arg 850 855 860	2592
TTA GGG AAT TTG AAA AAT GGA GTA AAA GAC ATT CAA AAG CAC AAA TGG Leu Gly Asn Leu Lys Asn Gly Val Lys Asp Ile Gln Lys His Lys Trp 865 870 875 880	2640
TTT GAG GGC TTT AAC TGG GAA GGC TTA AGA AAA GGT ACC TTG ACA CCT Phe Glu Gly Phe Asn Trp Glu Gly Leu Arg Lys Gly Thr Leu Thr Pro 885 890 895	2688
CCT ATA ATA CCA AGT GTT GCA TCA CCC ACA GAC ACA AGT AAT TTT GAC Pro Ile Ile Pro Ser Val Ala Ser Pro Thr Asp Thr Ser Asn Phe Asp 900 905 910	2736
AGT TTC CCT GAG GAC AAC GAT GAA CCA CCA CCT GAT GAC AAC TCA GGA Ser Phe Pro Glu Asp Asn Asp Glu Pro Pro Pro Asp Asp Asn Ser Gly 915 920 925	2784
TGG GAT ATA GAC TTC TAA Trp Asp Ile Asp Phe 930	2802

(2) INFORMATION FOR SEQ ID NO:135:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 933 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu 1 5 10 15
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly 20 25 30
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile 35 40 45
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr 50 55 60
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys 65 70 75 80
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu 85 90 95
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu 100 105 110
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly

115	120	125
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr		
130	135	140
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn		
145	150	155
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser		
165	170	175
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly		
180	185	190
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu		
195	200	205
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe		
210	215	220
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser		
225	230	235
Gly Leu Arg Ser Arg Gly Ser Met Gly Thr Leu Arg Asp Leu Gln Tyr		
245	250	255
Ala Leu Gln Glu Lys Ile Glu Glu Leu Arg Gln Arg Asp Ala Leu Ile		
260	265	270
Asp Glu Leu Glu Leu Glu Leu Asp Gln Lys Asp Glu Leu Ile Gln Lys		
275	280	285
Leu Gln Asn Glu Leu Asp Lys Tyr Arg Ser Val Ile Arg Pro Ala Thr		
290	295	300
Gln Gln Ala Gln Lys Gln Ser Ala Ser Thr Leu Gln Gly Glu Pro Arg		
305	310	315
Thr Lys Arg Gln Ala Ile Ser Ala Glu Pro Thr Ala Phe Asp Ile Gln		
325	330	335
Asp Leu Ser His Val Thr Leu Pro Phe Tyr Pro Lys Ser Pro Gln Ser		
340	345	350
Lys Asp Leu Ile Lys Glu Ala Ile Leu Asp Asn Asp Phe Met Lys Asn		
355	360	365
Leu Glu Leu Ser Gln Ile Gln Glu Ile Val Asp Cys Met Tyr Pro Val		
370	375	380
Glu Tyr Gly Lys Asp Ser Cys Ile Ile Lys Glu Gly Asp Val Gly Ser		
385	390	395
Leu Val Tyr Val Met Glu Asp Gly Lys Val Glu Val Thr Lys Glu Gly		
405	410	415
Val Lys Leu Cys Thr Met Gly Pro Gly Lys Val Phe Gly Glu Leu Ala		
420	425	430
Ile Leu Tyr Asn Cys Thr Arg Thr Ala Thr Val Lys Thr Leu Val Asn		
435	440	445
Val Lys Leu Trp Ala Ile Asp Arg Gln Cys Phe Gln Thr Ile Met Met		
450	455	460
Arg Thr Gly Leu Ile Lys His Thr Glu Tyr Met Glu Phe Leu Lys Ser		
465	470	475
Val Pro Thr Phe Gln Ser Leu Pro Glu Glu Ile Leu Ser Lys Leu Ala		
485	490	495
Asp Val Leu Glu Glu Thr His Tyr Glu Asn Gly Glu Tyr Ile Ile Arg		
500	505	510
Gln Gly Ala Arg Gly Asp Thr Phe Phe Ile Ile Ser Lys Gly Thr Val		
515	520	525
Asn Val Thr Arg Glu Asp Ser Pro Ser Glu Asp Pro Val Phe Leu Arg		
530	535	540
Thr Leu Gly Lys Gly Asp Trp Phe Gly Glu Lys Ala Leu Gln Gly Glu		
545	550	555
Asp Val Arg Thr Ala Asn Val Ile Ala Ala Glu Ala Val Thr Cys Leu		
565	570	575
Val Ile Asp Arg Asp Ser Phe Lys His Leu Ile Gly Gly Leu Asp Asp		

	580		585		590
Val	Ser	Asn	Lys	Ala	Tyr
	595		600		605
Ala	Glu	Ala	Ala	Phe	Phe
	610		615		620
Ile	Asp	Thr	Leu	Gly	Val
625			630		635
Leu	Lys	Ser	Glu	Glu	Ser
	645		650		655
Arg	His	Ile	Val	Asp	Thr
	660		665		670
Gln	Ile	Met	Gln	Gly	Ala
	675		680		685
Thr	Phe	Lys	Asp	Ser	Lys
	690		695		700
Gly	Gly	Glu	Leu	Trp	Thr
705			710		715
Ser	Thr	Thr	Arg	Phe	Tyr
	725		730		735
Leu	His	Ser	Lys	Gly	Ile
	740		745		750
Ile	Leu	Asp	His	Arg	Gly
	755		760		765
Lys	Lys	Ile	Gly	Phe	Gly
	770		775		780
Glu	Tyr	Val	Ala	Pro	Glu
785			790		795
Ala	Asp	Tyr	Trp	Ser	Leu
	805		810		815
Ser	Pro	Pro	Phe	Ser	Gly
	820		825		830
Leu	Arg	Gly	Ile	Asp	Met
	835		840		845
Ala	Ala	Asn	Leu	Ile	Lys
	850		855		860
Leu	Gly	Asn	Leu	Lys	Asn
865			870		875
Phe	Glu	Gly	Phe	Asn	Trp
	885		890		895
Pro	Ile	Ile	Pro	Ser	Val
	900		905		910
Ser	Phe	Pro	Glu	Asp	Asn
	915		920		925
Trp	Asp	Ile	Asp	Phe	
	930				

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2799 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence

(B) LOCATION: 1...2795

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

ATG GGC ACC TTG CCG GAT TTA CAG TAC GCG CTC CAG GAG AAG ATC GAG	48
Met Gly Thr Leu Arg Asp Leu Gln Tyr Ala Leu Gln Glu Lys Ile Glu	
1 5 10 15	
GAG CTG AGG CAG CCG GAT GCT CTC ATC GAC GAG CTG GAG CTG GAG TTG	96
Glu Leu Arg Gln Arg Asp Ala Leu Ile Asp Glu Leu Glu Leu Glu Leu	
20 25 30	
GAT CAG AAG GAC GAA CTG ATC CAG AAG CTG CAG AAC GAG CTG GAC AAG	144
Asp Gln Lys Asp Glu Leu Ile Gln Lys Leu Gln Asn Glu Leu Asp Lys	
35 40 45	
TAC GCG TCG GTG ATC CGA CCA GCC ACC CAG CAG GCG CAG AAG CAG AGC	192
Tyr Arg Ser Val Ile Arg Pro Ala Thr Gln Gln Ala Gln Lys Gln Ser	
50 55 60	
GCG AGC ACC TTG CAG GGC GAG CCG CGC ACC AAG CGG CAG GCG ATC TCC	240
Ala Ser Thr Leu Gln Gly Glu Pro Arg Thr Lys Arg Gln Ala Ile Ser	
65 70 75 80	
GCC GAG CCC ACC GCC TTC GAC ATC CAG GAT CTC AGC CAT GTG ACC CTG	288
Ala Glu Pro Thr Ala Phe Asp Ile Gln Asp Leu Ser His Val Thr Leu	
85 90 95	
CCC TTC TAC CCC AAG AGC CCA CAG TCC AAG GAT CTT ATA AAG GAA GCT	336
Pro Phe Tyr Pro Lys Ser Pro Gln Ser Lys Asp Leu Ile Lys Glu Ala	
100 105 110	
ATC CTT GAC AAT GAC TTT ATG AAG AAC TTG GAG CTG TCG CAG ATC CAG	384
Ile Leu Asp Asn Asp Phe Met Lys Asn Leu Glu Leu Ser Gln Ile Gln	
115 120 125	
GAG ATT GTG GAT TGT ATG TAC CCG GTG GAG TAT GGC AAG GAC AGT TGC	432
Glu Ile Val Asp Cys Met Tyr Pro Val Glu Tyr Gly Lys Asp Ser Cys	
130 135 140	
ATC ATC AAA GAA GGA GAC GTG GGG TCA CTG GTG TAT GTC ATG GAA GAT	480
Ile Ile Lys Glu Gly Asp Val Gly Ser Leu Val Tyr Val Met Glu Asp	
145 150 155 160	
GGT AAG GTT GAA GTT ACA AAA GAA GGT GTG AAG TTG TGT ACC ATG GGT	528
Gly Lys Val Glu Val Thr Lys Glu Gly Val Lys Leu Cys Thr Met Gly	
165 170 175	
CCA GGA AAA GTG TTT GGG GAA TTG GCT ATT CTT TAC AAC TGT ACC CGG	576
Pro Gly Lys Val Phe Gly Glu Leu Ala Ile Leu Tyr Asn Cys Thr Arg	
180 185 190	
ACA GCG ACC GTC AAG ACT CTT GTA AAT GTA AAA CTC TGG GCC ATT GAT	624
Thr Ala Thr Val Lys Thr Leu Val Asn Val Lys Leu Trp Ala Ile Asp	
195 200 205	
CGA CAA TGT TTT CAA ACA ATA ATG ATG AGG ACA GGA CTC ATC AAG CAT	672

Arg	Gln	Cys	Phe	Gln	Thr	Ile	Met	Met	Arg	Thr	Gly	Leu	Ile	Lys	His	
210						215					220					
ACC	GAG	TAT	ATG	GAA	TTT	TTA	AAA	AGC	GTT	CCA	ACA	TTC	CAG	AGC	CTT	720
Thr	Glu	Tyr	Met	Glu	Phe	Leu	Lys	Ser	Val	Pro	Thr	Phe	Gln	Ser	Leu	
225					230				235						240	
CCT	GAA	GAG	ATC	CTC	AGC	AAG	CTT	GCT	GAT	GTC	CTT	GAA	GAG	ACC	CAC	768
Pro	Glu	Glu	Ile	Leu	Ser	Lys	Leu	Ala	Asp	Val	Leu	Glu	Glu	Thr	His	
				245					250					255		
TAT	GAA	AAT	GGA	GAA	TAT	ATT	ATC	AGG	CAA	GGT	GCA	AGA	GGG	GAC	ACC	816
Tyr	Glu	Asn	Gly	Glu	Tyr	Ile	Ile	Arg	Gln	Gly	Ala	Arg	Gly	Asp	Thr	
			260					265					270			
TTC	TTT	ATC	ATC	AGC	AAA	GGA	ACG	GTA	AAT	GTC	ACT	CGT	GAA	GAC	TCA	864
Phe	Phe	Ile	Ile	Ser	Lys	Gly	Thr	Val	Asn	Val	Thr	Arg	Glu	Asp	Ser	
			275					280					285			
CCG	AGT	GAA	GAC	CCA	GTC	TTT	CTT	AGA	ACT	TTA	GGA	AAA	GGA	GAC	TGG	912
Pro	Ser	Glu	Asp	Pro	Val	Phe	Leu	Arg	Thr	Leu	Gly	Lys	Gly	Asp	Trp	
	290					295					300					
TTT	GGA	GAG	AAA	GCC	TTG	CAG	GGG	GAA	GAT	GTG	AGA	ACA	GCA	AAC	GTA	960
Phe	Gly	Glu	Lys	Ala	Leu	Gln	Gly	Glu	Asp	Val	Arg	Thr	Ala	Asn	Val	
305					310				315						320	
ATT	GCT	GCA	GAA	GCT	GTA	ACC	TGC	CTT	GTG	ATT	GAC	AGA	GAC	TCT	TTT	1008
Ile	Ala	Ala	Glu	Ala	Val	Thr	Cys	Leu	Val	Ile	Asp	Arg	Asp	Ser	Phe	
				325					330					335		
AAA	CAT	TTG	ATT	GGA	GGG	CTG	GAT	GAT	GTT	TCT	AAT	AAA	GCA	TAT	GAA	1056
Lys	His	Leu	Ile	Gly	Gly	Leu	Asp	Asp	Val	Ser	Asn	Lys	Ala	Tyr	Glu	
			340					345					350			
GAT	GCA	GAA	GCT	AAA	GCA	AAA	TAT	GAA	GCT	GAA	GCG	GCT	TTC	TTC	GCC	1104
Asp	Ala	Glu	Ala	Lys	Ala	Lys	Tyr	Glu	Ala	Glu	Ala	Ala	Phe	Phe	Ala	
			355					360					365			
AAC	CTG	AAG	CTG	TCT	GAT	TTC	AAC	ATC	ATT	GAT	ACC	CTT	GGA	GTT	GGA	1152
Asn	Leu	Lys	Leu	Ser	Asp	Phe	Asn	Ile	Ile	Asp	Thr	Leu	Gly	Val	Gly	
			370				375				380					
GGT	TTC	GGA	CGA	GTA	GAA	CTG	GTC	CAG	TTG	AAA	AGT	GAA	GAA	TCC	AAA	1200
Gly	Phe	Gly	Arg	Val	Glu	Leu	Val	Gln	Leu	Lys	Ser	Glu	Glu	Ser	Lys	
385					390				395						400	
ACG	TTT	GCA	ATG	AAG	ATT	CTC	AAG	AAA	CGT	CAC	ATT	GTG	GAC	ACA	AGA	1248
Thr	Phe	Ala	Met	Lys	Ile	Leu	Lys	Lys	Arg	His	Ile	Val	Asp	Thr	Arg	
				405					410					415		
CAG	CAG	GAG	CAC	ATC	CGC	TCA	GAG	AAG	CAG	ATC	ATG	CAG	GGG	GCT	CAT	1296
Gln	Gln	Glu	His	Ile	Arg	Ser	Glu	Lys	Gln	Ile	Met	Gln	Gly	Ala	His	
			420					425					430			
TCC	GAT	TTC	ATA	GTG	AGA	CTG	TAC	AGA	ACA	TTT	AAG	GAC	AGC	AAA	TAT	1344
Ser	Asp	Phe	Ile	Val	Arg	Leu	Tyr	Arg	Thr	Phe	Lys	Asp	Ser	Lys	Tyr	
			435				440					445				

TTG TAT ATG TTG ATG GAA GCT TGT CTA GGT GGA GAG CTC TGG ACC ATT	1392
Leu Tyr Met Leu Met Glu Ala Cys Leu Gly Gly Glu Leu Trp Thr Ile	
450 455 460	
CTC AGG GAT AGA GGT TCG TTT GAA GAT TCT ACA ACC AGA TTT TAC ACA	1440
Leu Arg Asp Arg Gly Ser Phe Glu Asp Ser Thr Thr Arg Phe Tyr Thr	
465 470 475 480	
GCA TGT GTG GTA GAA GCT TTT GCC TAT CTG CAT TCC AAA GGA ATC ATT	1488
Ala Cys Val Val Glu Ala Phe Ala Tyr Leu His Ser Lys Gly Ile Ile	
485 490 495	
TAC AGG GAC CTC AAG CCA GAA AAT CTC ATC CTA GAT CAC CGA GGT TAT	1536
Tyr Arg Asp Leu Lys Pro Glu Asn Leu Ile Leu Asp His Arg Gly Tyr	
500 505 510	
GCC AAA CTG GTT GAT TTT GGC TTT GCA AAG AAA ATA GGA TTT GGA AAG	1584
Ala Lys Leu Val Asp Phe Gly Phe Ala Lys Lys Ile Gly Phe Gly Lys	
515 520 525	
AAA ACA TGG ACT TTT TGT GGG ACT CCA GAG TAT GTA GCC CCA GAG ATC	1632
Lys Thr Trp Thr Phe Cys Gly Thr Pro Glu Tyr Val Ala Pro Glu Ile	
530 535 540	
ATC CTG AAC AAA GGC CAT GAC ATT TCA GCC GAC TAC TGG TCA CTG GGA	1680
Ile Leu Asn Lys Gly His Asp Ile Ser Ala Asp Tyr Trp Ser Leu Gly	
545 550 555 560	
ATC CTA ATG TAT GAA CTC CTG ACT GGC AGC CCA CCT TTC TCA GGC CCA	1728
Ile Leu Met Tyr Glu Leu Leu Thr Gly Ser Pro Pro Phe Ser Gly Pro	
565 570 575	
GAT CCT ATG AAA ACC TAT AAC ATC ATA TTG AGG GGG ATT GAC ATG ATA	1776
Asp Pro Met Lys Thr Tyr Asn Ile Ile Leu Arg Gly Ile Asp Met Ile	
580 585 590	
GAA TTT CCA AAG AAG ATT GCC AAA AAT GCT GCT AAT TTA ATT AAA AAA	1824
Glu Phe Pro Lys Lys Ile Ala Lys Asn Ala Ala Asn Leu Ile Lys Lys	
595 600 605	
CTA TGC AGG GAC AAT CCA TCA GAA AGA TTA GGG AAT TTG AAA AAT GGA	1872
Leu Cys Arg Asp Asn Pro Ser Glu Arg Leu Gly Asn Leu Lys Asn Gly	
610 615 620	
GTA AAA GAC ATT CAA AAG CAC AAA TGG TTT GAG GGC TTT AAC TGG GAA	1920
Val Lys Asp Ile Gln Lys His Lys Trp Phe Glu Gly Phe Asn Trp Glu	
625 630 635 640	
GGC TTA AGA AAA GGT ACC TTG ACA CCT CCT ATA ATA CCA AGT GTT GCA	1968
Gly Leu Arg Lys Gly Thr Leu Thr Pro Pro Ile Ile Pro Ser Val Ala	
645 650 655	
TCA CCC ACA GAC ACA AGT AAT TTT GAC AGT TTC CCT GAG GAC AAC GAT	2016
Ser Pro Thr Asp Thr Ser Asn Phe Asp Ser Phe Pro Glu Asp Asn Asp	
660 665 670	
GAA CCA CCA CCT GAT GAC AAC TCA GGA TGG GAT ATA GAC TTC TCG GAT	2064

Glu	Pro	Pro	Pro	Asp	Asp	Asn	Ser	Gly	Trp	Asp	Ile	Asp	Phe	Ser	Asp		
		675					680					685					
CCA	CCG	GTC	GCC	ACC	ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	2112	
Pro	Pro	Val	Ala	Thr	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly		
		690					695					700					
GTG	GTG	CCC	ATC	CTG	GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	2160	
Val	Val	Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys		
		705				710				715					720		
TTC	AGC	GTG	TCC	GGC	GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	2208	
Phe	Ser	Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu		
				725					730					735			
ACC	CTG	AAG	TTC	ATC	TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	2256	
Thr	Leu	Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro		
			740					745					750				
ACC	CTC	GTG	ACC	ACC	CTG	ACC	TAC	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	2304	
Thr	Leu	Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr		
		755					760					765					
CCC	GAC	CAC	ATG	AAG	CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	2352	
Pro	Asp	His	Met	Lys	Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu		
		770				775					780						
GGC	TAC	GTC	CAG	GAG	CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	2400	
Gly	Tyr	Val	Gln	Glu	Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr		
					790					795					800		
AAG	ACC	CGC	GCC	GAG	GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	2448	
Lys	Thr	Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg		
			805					810						815			
ATC	GAG	CTG	AAG	GGC	ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	2496	
Ile	Glu	Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly		
			820				825						830				
CAC	AAG	CTG	GAG	TAC	AAC	TAC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	2544	
His	Lys	Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala		
		835					840					845					
GAC	AAG	CAG	AAG	AAC	GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	2592	
Asp	Lys	Gln	Lys	Asn	Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn		
		850				855					860						
ATC	GAG	GAC	GGC	AGC	GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	2640	
Ile	Glu	Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr		
					865		870			875				880			
CCC	ATC	GGC	GAC	GGC	CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	2688	
Pro	Ile	Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser		
				885					890					895			
ACC	CAG	TCC	GCC	CTG	AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	2736	
Thr	Gln	Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met		
			900					905						910			

GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC 2784
Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp
915 920 925

GAG CTG TAC AA GTAA 2799
Glu Leu Tyr Lys
930

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 932 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

Met	Gly	Thr	Leu	Arg	Asp	Leu	Gln	Tyr	Ala	Leu	Gln	Glu	Lys	Ile	Glu
1				5					10					15	
Glu	Leu	Arg	Gln	Arg	Asp	Ala	Leu	Ile	Asp	Glu	Leu	Glu	Leu	Glu	Leu
			20					25					30		
Asp	Gln	Lys	Asp	Glu	Leu	Ile	Gln	Lys	Leu	Gln	Asn	Glu	Leu	Asp	Lys
		35				40					45				
Tyr	Arg	Ser	Val	Ile	Arg	Pro	Ala	Thr	Gln	Gln	Ala	Gln	Lys	Gln	Ser
	50					55					60				
Ala	Ser	Thr	Leu	Gln	Gly	Glu	Pro	Arg	Thr	Lys	Arg	Gln	Ala	Ile	Ser
65				70						75				80	
Ala	Glu	Pro	Thr	Ala	Phe	Asp	Ile	Gln	Asp	Leu	Ser	His	Val	Thr	Leu
				85					90					95	
Pro	Phe	Tyr	Pro	Lys	Ser	Pro	Gln	Ser	Lys	Asp	Leu	Ile	Lys	Glu	Ala
			100					105					110		
Ile	Leu	Asp	Asn	Asp	Phe	Met	Lys	Asn	Leu	Glu	Leu	Ser	Gln	Ile	Gln
		115					120					125			
Glu	Ile	Val	Asp	Cys	Met	Tyr	Pro	Val	Glu	Tyr	Gly	Lys	Asp	Ser	Cys
	130					135					140				
Ile	Ile	Lys	Glu	Gly	Asp	Val	Gly	Ser	Leu	Val	Tyr	Val	Met	Glu	Asp
145				150						155				160	
Gly	Lys	Val	Glu	Val	Thr	Lys	Glu	Gly	Val	Lys	Leu	Cys	Thr	Met	Gly
				165					170					175	
Pro	Gly	Lys	Val	Phe	Gly	Glu	Leu	Ala	Ile	Leu	Tyr	Asn	Cys	Thr	Arg
			180					185					190		
Thr	Ala	Thr	Val	Lys	Thr	Leu	Val	Asn	Val	Lys	Leu	Trp	Ala	Ile	Asp
		195				200						205			
Arg	Gln	Cys	Phe	Gln	Thr	Ile	Met	Met	Arg	Thr	Gly	Leu	Ile	Lys	His
	210					215					220				
Thr	Glu	Tyr	Met	Glu	Phe	Leu	Lys	Ser	Val	Pro	Thr	Phe	Gln	Ser	Leu
225					230					235					240
Pro	Glu	Glu	Ile	Leu	Ser	Lys	Leu	Ala	Asp	Val	Leu	Glu	Glu	Thr	His
				245					250					255	
Tyr	Glu	Asn	Gly	Glu	Tyr	Ile	Ile	Arg	Gln	Gly	Ala	Arg	Gly	Asp	Thr
		260						265					270		
Phe	Phe	Ile	Ile	Ser	Lys	Gly	Thr	Val	Asn	Val	Thr	Arg	Glu	Asp	Ser

275	280	285
Pro Ser Glu Asp	Pro Val Phe Leu Arg Thr	Leu Gly Lys Gly Asp Trp
290	295	300
Phe Gly Glu Lys	Ala Leu Gln Gly Glu Asp	Val Arg Thr Ala Asn Val
305	310	315
Ile Ala Ala Glu	Ala Val Thr Cys Leu Val	Ile Asp Arg Asp Ser Phe
	325	330
		335
Lys His Leu Ile	Gly Gly Leu Asp Asp	Val Ser Asn Lys Ala Tyr Glu
	340	345
		350
Asp Ala Glu Ala	Lys Ala Lys Tyr Glu Ala	Glu Ala Ala Phe Phe Ala
	355	360
		365
Asn Leu Lys Leu	Ser Asp Phe Asn Ile Ile	Asp Thr Leu Gly Val Gly
	370	375
		380
Gly Phe Gly Arg	Val Glu Leu Val Gln Leu	Lys Ser Glu Glu Ser Lys
385	390	395
Thr Phe Ala Met	Lys Ile Leu Lys Lys Arg	His Ile Val Asp Thr Arg
	405	410
		415
Gln Gln Glu His	Ile Arg Ser Glu Lys Gln	Ile Met Gln Gly Ala His
	420	425
		430
Ser Asp Phe Ile	Val Arg Leu Tyr Arg Thr	Phe Lys Asp Ser Lys Tyr
	435	440
		445
Leu Tyr Met Leu	Met Glu Ala Cys Leu Gly	Gly Glu Leu Trp Thr Ile
	450	455
		460
Leu Arg Asp Arg	Gly Ser Phe Glu Asp Ser	Thr Thr Arg Phe Tyr Thr
465	470	475
		480
Ala Cys Val Val	Glu Ala Phe Ala Tyr Leu	His Ser Lys Gly Ile Ile
	485	490
		495
Tyr Arg Asp Leu	Lys Pro Glu Asn Leu Ile	Leu Asp His Arg Gly Tyr
	500	505
		510
Ala Lys Leu Val	Asp Phe Gly Phe Ala Lys	Lys Ile Gly Phe Gly Lys
	515	520
		525
Lys Thr Trp Thr	Phe Cys Gly Thr Pro Glu	Tyr Val Ala Pro Glu Ile
	530	535
		540
Ile Leu Asn Lys	Gly His Asp Ile Ser Ala	Asp Tyr Trp Ser Leu Gly
545	550	555
		560
Ile Leu Met Tyr	Glu Leu Leu Thr Gly Ser	Pro Pro Phe Ser Gly Pro
	565	570
		575
Asp Pro Met Lys	Thr Tyr Asn Ile Ile Leu	Arg Gly Ile Asp Met Ile
	580	585
		590
Glu Phe Pro Lys	Lys Ile Ala Lys Asn Ala	Ala Asn Leu Ile Lys Lys
	595	600
		605
Leu Cys Arg Asp	Asn Pro Ser Glu Arg Leu	Gly Asn Leu Lys Asn Gly
	610	615
		620
Val Lys Asp Ile	Gln Lys His Lys Trp Phe	Glu Gly Phe Asn Trp Glu
625	630	635
		640
Gly Leu Arg Lys	Gly Thr Leu Thr Pro Pro	Ile Ile Pro Ser Val Ala
	645	650
		655
Ser Pro Thr Asp	Thr Ser Asn Phe Asp Ser	Phe Pro Glu Asp Asn Asp
	660	665
		670
Glu Pro Pro Pro	Asp Asp Asn Ser Gly Trp	Asp Ile Asp Phe Ser Asp
	675	680
		685
Pro Pro Val Ala	Thr Met Val Ser Lys Gly	Glu Glu Leu Phe Thr Gly
	690	695
		700
Val Val Pro Ile	Leu Val Glu Leu Asp Gly	Asp Val Asn Gly His Lys
705	710	715
		720
Phe Ser Val Ser	Gly Glu Gly Glu Gly Asp	Ala Thr Tyr Gly Lys Leu
	725	730
		735
Thr Leu Lys Phe	Ile Cys Thr Thr Gly Lys	Leu Pro Val Pro Trp Pro

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2184 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...2181
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	48
Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
1				5					10					15		
GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	96
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	
			20					25					30			
GAG	GGC	GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	AAG	TTC	ATC	144
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
		35					40					45				
TGC	ACC	ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	192
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
	50					55					60					

CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys 65 70 75 80	240
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu 85 90 95	288
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu 100 105 110	336
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly 115 120 125	384
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr 130 135 140	432
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn 145 150 155 160	480
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser 165 170 175	528
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly 180 185 190	576
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu 195 200 205	624
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe 210 215 220	672
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TCC Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser 225 230 235 240	720
GGA CTC AGA TCT CGA GGC ACC ATG AGC GAC GTG GCT ATT GTG AAG GAG Gly Leu Arg Ser Arg Gly Thr Met Ser Asp Val Ala Ile Val Lys Glu 245 250 255	768
GGT TGG CTG CAC AAA CGA GGG GAG TAC ATC AAG ACC TGG CGG CCA CGC Gly Trp Leu His Lys Arg Gly Glu Tyr Ile Lys Thr Trp Arg Pro Arg 260 265 270	816
TAC TTC CTC CTC AAG AAT GAT GGC ACC TTC ATT GGC TAC AAG GAG CGG Tyr Phe Leu Leu Lys Asn Asp Gly Thr Phe Ile Gly Tyr Lys Glu Arg 275 280 285	864
CCG CAG GAT GTG GAC CAA CGT GAG GCT CCC CTC AAC AAC TTC TCT GTG	912

Pro Gln Asp Val Asp Gln Arg Glu Ala Pro Leu Asn Asn Phe Ser Val	
290 295 300	
GCG CAG TGC CAG CTG ATG AAG ACG GAG CGG CCC CGG CCC AAC ACC TTC	960
Ala Gln Cys Gln Leu Met Lys Thr Glu Arg Pro Arg Pro Asn Thr Phe	
305 310 315 320	
ATC ATC CGC TGC CTG CAG TGG ACC ACT GTC ATC GAA CGC ACC TTC CAT	1008
Ile Ile Arg Cys Leu Gln Trp Thr Thr Val Ile Glu Arg Thr Phe His	
325 330 335	
GTG GAG ACT CCT GAG GAG CGG GAG GAG TGG ACA ACC GCC ATC CAG ACT	1056
Val Glu Thr Pro Glu Glu Arg Glu Glu Trp Thr Thr Ala Ile Gln Thr	
340 345 350	
GTG GCT GAC GGC CTC AAG AAG CAG GAG GAG GAG GAG ATG GAC TTC CGG	1104
Val Ala Asp Gly Leu Lys Lys Gln Glu Glu Glu Glu Met Asp Phe Arg	
355 360 365	
TCG GGC TCA CCC AGT GAC AAC TCA GGG GCT GAA GAG ATG GAG GTG TCC	1152
Ser Gly Ser Pro Ser Asp Asn Ser Ser Gly Ala Glu Glu Met Glu Val Ser	
370 375 380	
CTG GCC AAG CCC AAG CAC CGC GTG ACC ATG AAC GAG TTT GAG TAC CTG	1200
Leu Ala Lys Pro Lys His Arg Val Thr Met Asn Glu Phe Glu Tyr Leu	
385 390 395 400	
AAG CTG CTG GGC AAG GGC ACT TTC GGC AAG GTG ATC CTG GTG AAG GAG	1248
Lys Leu Leu Gly Lys Gly Thr Phe Gly Lys Val Ile Leu Val Lys Glu	
405 410 415	
AAG GCC ACA GGC CGC TAC TAC GCC ATG AAG ATC CTC AAG AAG GAA GTC	1296
Lys Ala Thr Gly Arg Tyr Tyr Ala Met Lys Ile Leu Lys Lys Glu Val	
420 425 430	
ATC GTG GCC AAG GAC GAG GTG GCC CAC ACA CTC ACC GAG AAC CGC GTC	1344
Ile Val Ala Lys Asp Glu Val Ala His Thr Leu Thr Glu Asn Arg Val	
435 440 445	
CTG CAG AAC TCC AGG CAC CCC TTC CTC ACA GCC CTG AAG TAC TCT TTC	1392
Leu Gln Asn Ser Arg His Pro Phe Leu Thr Ala Leu Lys Tyr Ser Phe	
450 455 460	
CAG ACC CAC GAC CGC CTC TGC TTT GTC ATG GAG TAC GCC AAC GGG GGC	1440
Gln Thr His Asp Arg Leu Cys Phe Val Met Glu Tyr Ala Asn Gly Gly	
465 470 475 480	
GAG CTG TTC TTC CAC CTG TCC CGG GAA CGT GTG TTC TCC GAG GAC CGG	1488
Glu Leu Phe Phe His Leu Ser Arg Glu Arg Val Phe Ser Glu Asp Arg	
485 490 495	
GCC CGC TTC TAT GGC GCT GAG ATT GTG TCA GCC CTG GAC TAC CTG CAC	1536
Ala Arg Phe Tyr Gly Ala Glu Ile Val Ser Ala Leu Asp Tyr Leu His	
500 505 510	
TCG GAG AAG AAC GTG GTG TAC CGG GAC CTC AAG CTG GAG AAC CTC ATG	1584
Ser Glu Lys Asn Val Val Tyr Arg Asp Leu Lys Leu Glu Asn Leu Met	
515 520 525	

CTG GAC AAG GAC GGG CAC ATT AAG ATC ACA GAC TTC GGG CTG TGC AAG	1632
Leu Asp Lys Asp Gly His Ile Lys Ile Thr Asp Phe Gly Leu Cys Lys	
530 535 540	
GAG GGG ATC AAG GAC GGT GCC ACC ATG AAG ACC TTT TGC GGC ACA CCT	1680
Glu Gly Ile Lys Asp Gly Ala Thr Met Lys Thr Phe Cys Gly Thr Pro	
545 550 555 560	
GAG TAC CTG GCC CCC GAG GTG CTG GAG GAC AAT GAC TAC GGC CGT GCA	1728
Glu Tyr Leu Ala Pro Glu Val Leu Glu Asp Asn Asp Tyr Gly Arg Ala	
565 570 575	
GTG GAC TGG TGG GGG CTG GGC GTG GTC ATG TAC GAG ATG ATG TGC GGT	1776
Val Asp Trp Trp Gly Leu Gly Val Val Met Tyr Glu Met Met Cys Gly	
580 585 590	
CGC CTG CCC TTC TAC AAC CAG GAC CAT GAG AAG CTT TTT GAG CTC ATC	1824
Arg Leu Pro Phe Tyr Asn Gln Asp His Glu Lys Leu Phe Glu Leu Ile	
595 600 605	
CTC ATG GAG GAG ATC CGC TTC CCG CGC ACG CTT GGT CCC GAG GCC AAG	1872
Leu Met Glu Glu Ile Arg Phe Pro Arg Thr Leu Gly Pro Glu Ala Lys	
610 615 620	
TCC TTG CTT TCA GGG CTG CTC AAG AAG GAC CCC AAG CAG AGG CTT GGC	1920
Ser Leu Leu Ser Gly Leu Leu Lys Lys Asp Pro Lys Gln Arg Leu Gly	
625 630 635 640	
GGG GGC TCC GAG GAC GCC AAG GAG ATC ATG CAG CAT CGC TTC TTT GCC	1968
Gly Gly Ser Glu Asp Ala Lys Glu Ile Met Gln His Arg Phe Phe Ala	
645 650 655	
GGT ATC GTG TGG CAG CAC GTG TAC GAG AAG AAG CTC AGC CCA CCC TTC	2016
Gly Ile Val Trp Gln His Val Tyr Glu Lys Lys Leu Ser Pro Pro Phe	
660 665 670	
AAG CCC CAG GTC ACG TCG GAG ACT GAC ACC AGG TAT TTT GAT GAG GAG	2064
Lys Pro Gln Val Thr Ser Glu Thr Asp Thr Arg Tyr Phe Asp Glu Glu	
675 680 685	
TTC ACG GCC CAG ATG ATC ACC ATC ACA CCA CCT GAC CAA GAT GAC AGC	2112
Phe Thr Ala Gln Met Ile Thr Ile Thr Pro Pro Asp Gln Asp Asp Ser	
690 695 700	
ATG GAG TGT GTG GAC AGC GAG CGC AGG CCC CAC TTC CCC CAG TTC TCC	2160
Met Glu Cys Val Asp Ser Glu Arg Arg Pro His Phe Pro Gln Phe Ser	
705 710 715 720	
TAC TCG GCC AGC AGC ACG GCC TGA	2184
Tyr Ser Ala Ser Ser Thr Ala	
725	

(2) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 727 amino acids

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu
1				5					10					15	
Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly
			20					25					30		
Glu	Gly	Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile
			35					40					45		
Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr
	50					55					60				
Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys
65					70					75					80
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu
				85					90					95	
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu
			100					105					110		
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly
		115					120					125			
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr
	130					135					140				
Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn
145					150					155					160
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser
			165					170					175		
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly
			180					185					190		
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu
		195				200						205			
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe
	210					215				220					
Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Ser
225					230					235					240
Gly	Leu	Arg	Ser	Arg	Gly	Thr	Met	Ser	Asp	Val	Ala	Ile	Val	Lys	Glu
			245					250					255		
Gly	Trp	Leu	His	Lys	Arg	Gly	Glu	Tyr	Ile	Lys	Thr	Trp	Arg	Pro	Arg
		260					265					270			
Tyr	Phe	Leu	Leu	Lys	Asn	Asp	Gly	Thr	Phe	Ile	Gly	Tyr	Lys	Glu	Arg
		275					280					285			
Pro	Gln	Asp	Val	Asp	Gln	Arg	Glu	Ala	Pro	Leu	Asn	Asn	Phe	Ser	Val
	290					295					300				
Ala	Gln	Cys	Gln	Leu	Met	Lys	Thr	Glu	Arg	Pro	Arg	Pro	Asn	Thr	Phe
305					310					315					320
Ile	Ile	Arg	Cys	Leu	Gln	Trp	Thr	Thr	Val	Ile	Glu	Arg	Thr	Phe	His
			325						330				335		
Val	Glu	Thr	Pro	Glu	Glu	Arg	Glu	Glu	Trp	Thr	Thr	Ala	Ile	Gln	Thr
			340				345					350			
Val	Ala	Asp	Gly	Leu	Lys	Lys	Gln	Glu	Glu	Glu	Glu	Met	Asp	Phe	Arg
		355					360					365			
Ser	Gly	Ser	Pro	Ser	Asp	Asn	Ser	Gly	Ala	Glu	Glu	Met	Glu	Val	Ser
	370					375					380				
Leu	Ala	Lys	Pro	Lys	His	Arg	Val	Thr	Met	Asn	Glu	Phe	Glu	Tyr	Leu

385		390		395		400									
Lys	Leu	Leu	Gly	Lys	Gly	Thr	Phe	Gly	Lys	Val	Ile	Leu	Val	Lys	Glu
		405						410						415	
Lys	Ala	Thr	Gly	Arg	Tyr	Tyr	Ala	Met	Lys	Ile	Leu	Lys	Lys	Glu	Val
		420						425						430	
Ile	Val	Ala	Lys	Asp	Glu	Val	Ala	His	Thr	Leu	Thr	Glu	Asn	Arg	Val
		435						440						445	
Leu	Gln	Asn	Ser	Arg	His	Pro	Phe	Leu	Thr	Ala	Leu	Lys	Tyr	Ser	Phe
		450						455						460	
Gln	Thr	His	Asp	Arg	Leu	Cys	Phe	Val	Met	Glu	Tyr	Ala	Asn	Gly	Gly
		465						470						475	
Glu	Leu	Phe	Phe	His	Leu	Ser	Arg	Glu	Arg	Val	Phe	Ser	Glu	Asp	Arg
			485						490					495	
Ala	Arg	Phe	Tyr	Gly	Ala	Glu	Ile	Val	Ser	Ala	Leu	Asp	Tyr	Leu	His
			500						505					510	
Ser	Glu	Lys	Asn	Val	Val	Tyr	Arg	Asp	Leu	Lys	Leu	Glu	Asn	Leu	Met
		515						520						525	
Leu	Asp	Lys	Asp	Gly	His	Ile	Lys	Ile	Thr	Asp	Phe	Gly	Leu	Cys	Lys
		530						535						540	
Glu	Gly	Ile	Lys	Asp	Gly	Ala	Thr	Met	Lys	Thr	Phe	Cys	Gly	Thr	Pro
		545						550						555	
Glu	Tyr	Leu	Ala	Pro	Glu	Val	Leu	Glu	Asp	Asn	Asp	Tyr	Gly	Arg	Ala
			565						570					575	
Val	Asp	Trp	Trp	Gly	Leu	Gly	Val	Val	Met	Tyr	Glu	Met	Met	Cys	Gly
			580						585					590	
Arg	Leu	Pro	Phe	Tyr	Asn	Gln	Asp	His	Glu	Lys	Leu	Phe	Glu	Leu	Ile
		595						600						605	
Leu	Met	Glu	Glu	Ile	Arg	Phe	Pro	Arg	Thr	Leu	Gly	Pro	Glu	Ala	Lys
		610						615						620	
Ser	Leu	Leu	Ser	Gly	Leu	Leu	Lys	Lys	Asp	Pro	Lys	Gln	Arg	Leu	Gly
		625						630						635	
Gly	Gly	Ser	Glu	Asp	Ala	Lys	Glu	Ile	Met	Gln	His	Arg	Phe	Phe	Ala
			645						650					655	
Gly	Ile	Val	Trp	Gln	His	Val	Tyr	Glu	Lys	Lys	Leu	Ser	Pro	Pro	Phe
			660						665					670	
Lys	Pro	Gln	Val	Thr	Ser	Glu	Thr	Asp	Thr	Arg	Tyr	Phe	Asp	Glu	Glu
			675					680						685	
Phe	Thr	Ala	Gln	Met	Ile	Thr	Ile	Thr	Pro	Pro	Asp	Gln	Asp	Asp	Ser
			690					695						700	
Met	Glu	Cys	Val	Asp	Ser	Glu	Arg	Arg	Pro	His	Phe	Pro	Gln	Phe	Ser
			705					710						715	
Tyr	Ser	Ala	Ser	Ser	Thr	Ala									720
															725

(2) INFORMATION FOR SEQ ID NO:140:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2394 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2391
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

ATG GAC GAA CTG TTC CCC CTC ATC TTC CCG GCA GAG CCA GCC CAG GCC	48
Met Asp Glu Leu Phe Pro Leu Ile Phe Pro Ala Glu Pro Ala Gln Ala	
1 5 10 15	
TCT GGC CCC TAT GTG GAG ATC ATT GAG CAG CCC AAG CAG CGG GGC ATG	96
Ser Gly Pro Tyr Val Glu Ile Ile Glu Gln Pro Lys Gln Arg Gly Met	
20 25 30	
CGC TTC CGC TAC AAG TGC GAG GGG CGC TCC GCG GGC AGC ATC CCA GGC	144
Arg Phe Arg Tyr Lys Cys Glu Gly Arg Ser Ala Gly Ser Ile Pro Gly	
35 40 45	
GAG AGG AGC ACA GAT ACC ACC AAG ACC CAC CCC ACC ATC AAG ATC AAT	192
Glu Arg Ser Thr Asp Thr Lys Thr His Pro Thr Ile Lys Ile Asn	
50 55 60	
GGC TAC ACA GGA CCA GGG ACA GTG CGC ATC TCC CTG GTC ACC AAG GAC	240
Gly Tyr Thr Gly Pro Gly Thr Val Arg Ile Ser Leu Val Thr Lys Asp	
65 70 75 80	
CCT CCT CAC CGG CCT CAC CCC CAC GAG CTT GTA GGA AAG GAC TGC CGG	288
Pro Pro His Arg Pro His Pro His Glu Leu Val Gly Lys Asp Cys Arg	
85 90 95	
GAT GGC TTC TAT GAG GCT GAG CTC TGC CCG GAC CGC TGC ATC CAC AGT	336
Asp Gly Phe Tyr Glu Ala Glu Leu Cys Pro Asp Arg Cys Ile His Ser	
100 105 110	
TTC CAG AAC CTG GGA ATC CAG TGT GTG AAG AAG CGG GAC CTG GAG CAG	384
Phe Gln Asn Leu Gly Ile Gln Cys Val Lys Lys Arg Asp Leu Glu Gln	
115 120 125	
GCT ATC AGT CAG CGC ATC CAG ACC AAC AAC AAC CCC TTC CAA GTT CCT	432
Ala Ile Ser Gln Arg Ile Gln Thr Asn Asn Asn Pro Phe Gln Val Pro	
130 135 140	
ATA GAA GAG CAG CGT GGG GAC TAC GAC CTG AAT GCT GTG CGG CTC TGC	480
Ile Glu Glu Gln Arg Gly Asp Tyr Asp Leu Asn Ala Val Arg Leu Cys	
145 150 155 160	
TTC CAG GTG ACA GTG CGG GAC CCA TCA GGC AGG CCC CTC CGC CTG CCG	528
Phe Gln Val Thr Val Arg Asp Pro Ser Gly Arg Pro Leu Arg Leu Pro	
165 170 175	
CCT GTC CTT CCT CAT CCC ATC TTT GAC AAT CGT GCC CCC AAC ACT GCC	576
Pro Val Leu Pro His Pro Ile Phe Asp Asn Arg Ala Pro Asn Thr Ala	
180 185 190	
GAG CTC AAG ATC TGC CGA GTG AAC CGA AAC TCT GGC AGC TGC CTC GGT	624
Glu Leu Lys Ile Cys Arg Val Asn Arg Asn Ser Gly Ser Cys Leu Gly	
195 200 205	
GGG GAT GAG ATC TTC CTA CTG TGT GAC AAG GTG CAG AAA GAG GAC ATT	672
Gly Asp Glu Ile Phe Leu Leu Cys Asp Lys Val Gln Lys Glu Asp Ile	
210 215 220	

GAG GTG TAT TTC ACG GGA CCA GGC TGG GAG GCC CGA GGC TCC TTT TCG	720
Glu Val Tyr Phe Thr Gly Pro Gly Trp Glu Ala Arg Gly Ser Phe Ser	
225 230 235 240	
CAA GCT GAT GTG CAC CGA CAA GTG GCC ATT GTG TTC CGG ACC CCT CCC	768
Gln Ala Asp Val His Arg Gln Val Ala Ile Val Phe Arg Thr Pro Pro	
245 250 255	
TAC GCA GAC CCC AGC CTG CAG GCT CCT GTG CGT GTC TCC ATG CAG CTG	816
Tyr Ala Asp Pro Ser Leu Gln Ala Pro Val Arg Val Ser Met Gln Leu	
260 265 270	
CGG CGG CCT TCC GAC CGG GAG CTC AGT GAG CCC ATG GAA TTC CAG TAC	864
Arg Arg Pro Ser Asp Arg Glu Leu Ser Glu Pro Met Glu Phe Gln Tyr	
275 280 285	
CTG CCA GAT ACA GAC GAT CGT CAC CGG ATT GAG GAG AAA CGT AAA AGG	912
Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu Glu Lys Arg Lys Arg	
290 295 300	
ACA TAT GAG ACC TTC AAG AGC ATC ATG AAG AAG AGT CCT TTC AGC GGA	960
Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly	
305 310 315 320	
CCC ACC GAC CCC CGG CCT CCA CCT CGA CGC ATT GCT GTG CCT TCC CGC	1008
Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg	
325 330 335	
AGC TCA GCT TCT GTC CCC AAG CCA GCA CCC CAG CCC TAT CCC TTT ACG	1056
Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr	
340 345 350	
TCA TCC CTG AGC ACC ATC AAC TAT GAT GAG TTT CCC ACC ATG GTG TTT	1104
Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe	
355 360 365	
CCT TCT GGG CAG ATC AGC CAG GCC TCG GCC TTG GCC CCG GCC CCT CCC	1152
Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro	
370 375 380	
CAA GTC CTG CCC CAG GCT CCA GCC CCT GCC CCT GCT CCA GCC ATG GTA	1200
Gln Val Leu Pro Gln Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val	
385 390 395 400	
TCA GCT CTG GCC CAG GCC CCA GCC CCT GTC CCA GTC CTA GCC CCA GGC	1248
Ser Ala Leu Ala Gln Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly	
405 410 415	
CCT CCT CAG GCT GTG GCC CCA CCT GCC CCC AAG CCC ACC CAG GCT GGG	1296
Pro Pro Gln Ala Val Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly	
420 425 430	
GAA GGA ACG CTG TCA GAG GCC CTG CTG CAG CTG CAG TTT GAT GAT GAA	1344
Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu	
435 440 445	
GAC CTG GGG GCC TTG CTT GGC AAC AGC ACA GAC CCA GCT GTG TTC ACA	1392

Asp	Leu	Gly	Ala	Leu	Leu	Gly	Asn	Ser	Thr	Asp	Pro	Ala	Val	Phe	Thr	
450						455				460						
GAC	CTG	GCA	TCC	GTC	GAC	AAC	TCC	GAG	TTT	CAG	CAG	CTG	CTG	AAC	CAG	1440
Asp	Leu	Ala	Ser	Val	Asp	Asn	Ser	Glu	Phe	Gln	Gln	Leu	Leu	Asn	Gln	
465				470				475						480		
GGC	ATA	CCT	GTG	GCC	CCC	CAC	ACA	ACT	GAG	CCC	ATG	CTG	ATG	GAG	TAC	1488
Gly	Ile	Pro	Val	Ala	Pro	His	Thr	Thr	Glu	Pro	Met	Leu	Met	Glu	Tyr	
				485				490						495		
CCT	GAG	GCT	ATA	ACT	CGC	CTA	GTG	ACA	GGG	GCC	CAG	AGG	CCC	CCC	GAC	1536
Pro	Glu	Ala	Ile	Thr	Arg	Leu	Val	Thr	Gly	Ala	Gln	Arg	Pro	Pro	Asp	
			500					505					510			
CCA	GCT	CCT	GCT	CCA	CTG	GGG	GCC	CCG	GGG	CTC	CCC	AAT	GGC	CTC	CTT	1584
Pro	Ala	Pro	Ala	Pro	Leu	Gly	Ala	Pro	Gly	Leu	Pro	Asn	Gly	Leu	Leu	
		515				520					525					
TCA	GGA	GAT	GAA	GAC	TTC	TCC	TCC	ATT	GCG	GAC	ATG	GAC	TTC	TCA	GCC	1632
Ser	Gly	Asp	Glu	Asp	Phe	Ser	Ser	Ile	Ala	Asp	Met	Asp	Phe	Ser	Ala	
	530					535					540					
CTG	CTG	AGT	CAG	ATC	AGC	TCC	TTG	GAT	CCA	CCG	GTC	GCC	ACC	ATG	GTG	1680
Leu	Leu	Ser	Gln	Ile	Ser	Ser	Leu	Asp	Pro	Pro	Val	Ala	Thr	Met	Val	
545				550						555				560		
AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	GTC	GAG	1728
Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	Val	Glu	
			565					570					575			
CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	GAG	GGC	1776
Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	Glu	Gly	
		580						585					590			
GAG	GGC	GAT	GCC	ACC	TAC	GGC	AAG	CTG	ACC	CTG	AAG	TTC	ATC	TGC	ACC	1824
Glu	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	Thr	
		595					600					605				
ACC	GGC	AAG	CTG	CCC	GTG	CCC	TGG	CCC	ACC	CTC	GTG	ACC	ACC	CTG	ACC	1872
Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	Leu	Thr	
	610					615					620					
TAC	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	CAC	ATG	AAG	CAG	CAC	1920
Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	Gln	His	
625				630						635				640		
GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	CGC	ACC	1968
Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	Arg	Thr	
			645					650					655			
ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	GTG	AAG	2016
Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val	Lys	
			660					665					670			
TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	CTG	AAG	GGC	ATC	GAC	2064
Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile	Asp	
		675					680					685				

TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC	2112
Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr	
690 695 700	
AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC	2160
Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile	
705 710 715 720	
AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG	2208
Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln	
725 730 735	
CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG	2256
Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val	
740 745 750	
CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA	2304
Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys	
755 760 765	
GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC	2352
Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr	
770 775 780	
GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA	2394
Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys	
785 790 795	

(2) INFORMATION FOR SEQ ID NO:141:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 797 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

Met Asp Glu Leu Phe Pro Leu Ile Phe Pro Ala Glu Pro Ala Gln Ala	
1 5 10 15	
Ser Gly Pro Tyr Val Glu Ile Ile Glu Gln Pro Lys Gln Arg Gly Met	
20 25 30	
Arg Phe Arg Tyr Lys Cys Glu Gly Arg Ser Ala Gly Ser Ile Pro Gly	
35 40 45	
Glu Arg Ser Thr Asp Thr Thr Lys Thr His Pro Thr Ile Lys Ile Asn	
50 55 60	
Gly Tyr Thr Gly Pro Gly Thr Val Arg Ile Ser Leu Val Thr Lys Asp	
65 70 75 80	
Pro Pro His Arg Pro His Pro His Glu Leu Val Gly Lys Asp Cys Arg	
85 90 95	
Asp Gly Phe Tyr Glu Ala Glu Leu Cys Pro Asp Arg Cys Ile His Ser	
100 105 110	
Phe Gln Asn Leu Gly Ile Gln Cys Val Lys Lys Arg Asp Leu Glu Gln	

115	120	125
Ala Ile Ser Gln Arg Ile Gln Thr Asn Asn Asn Pro Phe Gln Val Pro		
130	135	140
Ile Glu Glu Gln Arg Gly Asp Tyr Asp Leu Asn Ala Val Arg Leu Cys		
145	150	155
Phe Gln Val Thr Val Arg Asp Pro Ser Gly Arg Pro Leu Arg Leu Pro		
165	170	175
Pro Val Leu Pro His Pro Ile Phe Asp Asn Arg Ala Pro Asn Thr Ala		
180	185	190
Glu Leu Lys Ile Cys Arg Val Asn Arg Asn Ser Gly Ser Cys Leu Gly		
195	200	205
Gly Asp Glu Ile Phe Leu Leu Cys Asp Lys Val Gln Lys Glu Asp Ile		
210	215	220
Glu Val Tyr Phe Thr Gly Pro Gly Trp Glu Ala Arg Gly Ser Phe Ser		
225	230	235
Gln Ala Asp Val His Arg Gln Val Ala Ile Val Phe Arg Thr Pro Pro		
245	250	255
Tyr Ala Asp Pro Ser Leu Gln Ala Pro Val Arg Val Ser Met Gln Leu		
260	265	270
Arg Arg Pro Ser Asp Arg Glu Leu Ser Glu Pro Met Glu Phe Gln Tyr		
275	280	285
Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu Glu Lys Arg Lys Arg		
290	295	300
Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly		
305	310	315
Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg		
325	330	335
Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr		
340	345	350
Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe		
355	360	365
Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro		
370	375	380
Gln Val Leu Pro Gln Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val		
385	390	395
Ser Ala Leu Ala Gln Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly		
405	410	415
Pro Pro Gln Ala Val Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly		
420	425	430
Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu		
435	440	445
Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr		
450	455	460
Asp Leu Ala Ser Val Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln		
465	470	475
Gly Ile Pro Val Ala Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr		
485	490	495
Pro Glu Ala Ile Thr Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp		
500	505	510
Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu		
515	520	525
Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala		
530	535	540
Leu Leu Ser Gln Ile Ser Ser Leu Asp Pro Pro Val Ala Thr Met Val		
545	550	555
Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu		
565	570	575
Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly		

	580		585		590
Glu Gly Asp	Ala Thr Tyr Gly Lys	Leu Thr Leu Lys Phe	Ile Cys Thr		
595	600	605			
Thr Gly Lys	Leu Pro Val Pro Trp Pro Thr	Leu Val Thr Thr	Leu Thr		
610	615	620			
Tyr Gly Val	Gln Cys Phe Ser Arg Tyr Pro Asp	His Met Lys Gln His			
625	630	635	640		
Asp Phe Phe	Lys Ser Ala Met Pro Glu Gly Tyr Val	Gln Glu Arg Thr			
	645	650	655		
Ile Phe Phe	Lys Asp Asp Gly Asn Tyr Lys Thr Arg	Ala Glu Val Lys			
	660	665	670		
Phe Glu Gly	Asp Thr Leu Val Asn Arg Ile Glu Leu Lys	Gly Ile Asp			
	675	680	685		
Phe Lys Glu	Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr	Asn Tyr			
690	695	700			
Asn Ser His	Asn Val Tyr Ile Met Ala Asp Lys Gln Lys	Asn Gly Ile			
705	710	715	720		
Lys Val Asn	Phe Lys Ile Arg His Asn Ile Glu Asp Gly	Ser Val Gln			
	725	730	735		
Leu Ala Asp	His Tyr Gln Gln Asn Thr Pro Ile Gly Asp	Gly Pro Val			
	740	745	750		
Leu Leu Pro	Asp Asn His Tyr Leu Ser Thr Gln Ser Ala	Leu Ser Lys			
	755	760	765		
Asp Pro Asn	Glu Lys Arg Asp His Met Val Leu Leu Glu Phe	Val Thr			
	770	775	780		
Ala Ala Gly	Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys				
785	790	795			

(2) INFORMATION FOR SEQ ID NO:142:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2394 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2391
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG	48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu	
1 5 10 15	
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC	96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly	
20 25 30	
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC	144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile	
35 40 45	
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC	192

Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	
50						55						60				
CTG	ACC	TAC	GGC	GTG	CAG	TGC	TTC	AGC	CGC	TAC	CCC	GAC	CAC	ATG	AAG	240
Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	
65					70				75						80	
CAG	CAC	GAC	TTC	TTC	AAG	TCC	GCC	ATG	CCC	GAA	GGC	TAC	GTC	CAG	GAG	288
Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	
				85					90					95		
CGC	ACC	ATC	TTC	TTC	AAG	GAC	GAC	GGC	AAC	TAC	AAG	ACC	CGC	GCC	GAG	336
Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	
			100					105						110		
GTG	AAG	TTC	GAG	GGC	GAC	ACC	CTG	GTG	AAC	CGC	ATC	GAG	CTG	AAG	GGC	384
Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	
			115				120					125				
ATC	GAC	TTC	AAG	GAG	GAC	GGC	AAC	ATC	CTG	GGG	CAC	AAG	CTG	GAG	TAC	432
Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	
			130				135					140				
AAC	TAC	AAC	AGC	CAC	AAC	GTC	TAT	ATC	ATG	GCC	GAC	AAG	CAG	AAG	AAC	480
Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	
145					150					155					160	
GGC	ATC	AAG	GTG	AAC	TTC	AAG	ATC	CGC	CAC	AAC	ATC	GAG	GAC	GGC	AGC	528
Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	
				165					170					175		
GTG	CAG	CTC	GCC	GAC	CAC	TAC	CAG	CAG	AAC	ACC	CCC	ATC	GGC	GAC	GGC	576
Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	
			180					185					190			
CCC	GTG	CTG	CTG	CCC	GAC	AAC	CAC	TAC	CTG	AGC	ACC	CAG	TCC	GCC	CTG	624
Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	
			195				200					205				
AGC	AAA	GAC	CCC	AAC	GAG	AAG	CGC	GAT	CAC	ATG	GTC	CTG	CTG	GAG	TTC	672
Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	
			210				215				220					
GTG	ACC	GCC	GCC	GGG	ATC	ACT	CTC	GGC	ATG	GAC	GAG	CTG	TAC	AAG	TCC	720
Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Ser	
225					230				235						240	
GGA	CTC	AGA	TCT	CGA	GCC	ATG	GAC	GAA	CTG	TTC	CCC	CTC	ATC	TTC	CCG	768
Gly	Leu	Arg	Ser	Arg	Ala	Met	Asp	Glu	Leu	Phe	Pro	Leu	Ile	Phe	Pro	
				245					250					255		
GCA	GAG	CCA	GCC	CAG	GCC	TCT	GGC	CCC	TAT	GTG	GAG	ATC	ATT	GAG	CAG	816
Ala	Glu	Pro	Ala	Gln	Ala	Ser	Gly	Pro	Tyr	Val	Glu	Ile	Ile	Glu	Gln	
			260					265					270			
CCC	AAG	CAG	CGG	GGC	ATG	CGC	TTC	CGC	TAC	AAG	TGC	GAG	GGG	CGC	TCC	864
Pro	Lys	Gln	Arg	Gly	Met	Arg	Phe	Arg	Tyr	Lys	Cys	Glu	Gly	Arg	Ser	
			275				280						285			

GCG GGC AGC ATC CCA GGC GAG AGG AGC ACA GAT ACC ACC AAG ACC CAC	912
Ala Gly Ser Ile Pro Gly Glu Arg Ser Thr Asp Thr Thr Lys Thr His	
290 295 300	
CCC ACC ATC AAG ATC AAT GGC TAC ACA GGA CCA GGG ACA GTG CGC ATC	960
Pro Thr Ile Lys Ile Asn Gly Tyr Thr Gly Pro Gly Thr Val Arg Ile	
305 310 315 320	
TCC CTG GTC ACC AAG GAC CCT CCT CAC CGG CCT CAC CCC CAC GAG CTT	1008
Ser Leu Val Thr Lys Asp Pro Pro His Arg Pro His Pro His Glu Leu	
325 330 335	
GTA GGA AAG GAC TGC CGG GAT GGC TTC TAT GAG GCT GAG CTC TGC CCG	1056
Val Gly Lys Asp Cys Arg Asp Gly Phe Tyr Glu Ala Glu Leu Cys Pro	
340 345 350	
GAC CGC TGC ATC CAC AGT TTC CAG AAC CTG GGA ATC CAG TGT GTG AAG	1104
Asp Arg Cys Ile His Ser Phe Gln Asn Leu Gly Ile Gln Cys Val Lys	
355 360 365	
AAG CGG GAC CTG GAG CAG GCT ATC AGT CAG CGC ATC CAG ACC AAC AAC	1152
Lys Arg Asp Leu Glu Gln Ala Ile Ser Gln Arg Ile Gln Thr Asn Asn	
370 375 380	
AAC CCC TTC CAA GTT CCT ATA GAA GAG CAG CGT GGG GAC TAC GAC CTG	1200
Asn Pro Phe Gln Val Pro Ile Glu Glu Gln Arg Gly Asp Tyr Asp Leu	
385 390 395 400	
AAT GCT GTG CGG CTC TGC TTC CAG GTG ACA GTG CGG GAC CCA TCA GGC	1248
Asn Ala Val Arg Leu Cys Phe Gln Val Thr Val Arg Asp Pro Ser Gly	
405 410 415	
AGG CCC CTC CGC CTG CCG CCT GTC CTT CCT CAT CCC ATC TTT GAC AAT	1296
Arg Pro Leu Arg Leu Pro Pro Val Leu Pro His Pro Ile Phe Asp Asn	
420 425 430	
CGT GCC CCC AAC ACT GCC GAG CTC AAG ATC TGC CGA GTG AAC CGA AAC	1344
Arg Ala Pro Asn Thr Ala Glu Leu Lys Ile Cys Arg Val Asn Arg Asn	
435 440 445	
TCT GGC AGC TGC CTC GGT GGG GAT GAG ATC TTC CTA CTG TGT GAC AAG	1392
Ser Gly Ser Cys Leu Gly Gly Asp Glu Ile Phe Leu Leu Cys Asp Lys	
450 455 460	
GTG CAG AAA GAG GAC ATT GAG GTG TAT TTC ACG GGA CCA GGC TGG GAG	1440
Val Gln Lys Glu Asp Ile Glu Val Tyr Phe Thr Gly Pro Gly Trp Glu	
465 470 475 480	
GCC CGA GGC TCC TTT TCG CAA GCT GAT GTG CAC CGA CAA GTG GCC ATT	1488
Ala Arg Gly Ser Phe Ser Gln Ala Asp Val His Arg Gln Val Ala Ile	
485 490 495	
GTG TTC CGG ACC CCT CCC TAC GCA GAC CCC AGC CTG CAG GCT CCT GTG	1536
Val Phe Arg Thr Pro Pro Tyr Ala Asp Pro Ser Leu Gln Ala Pro Val	
500 505 510	
CGT GTC TCC ATG CAG CTG CGG CGG CCT TCC GAC CGG GAG CTC AGT GAG	1584

Arg Val Ser Met Gln Leu Arg Arg Pro Ser Asp Arg Glu Leu Ser Glu	
515 520 525	
CCC ATG GAA TTC CAG TAC CTG CCA GAT ACA GAC GAT CGT CAC CGG ATT	1632
Pro Met Glu Phe Gln Tyr Leu Pro Asp Thr Asp Asp Arg His Arg Ile	
530 535 540	
GAG GAG AAA CGT AAA AGG ACA TAT GAG ACC TTC AAG AGC ATC ATG AAG	1680
Glu Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys	
545 550 555 560	
AAG AGT CCT TTC AGC GGA CCC ACC GAC CCC CGG CCT CCA CCT CGA CGC	1728
Lys Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg	
565 570 575	
ATT GCT GTG CCT TCC CGC AGC TCA GCT TCT GTC CCC AAG CCA GCA CCC	1776
Ile Ala Val Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala Pro	
580 585 590	
CAG CCC TAT CCC TTT ACG TCA TCC CTG AGC ACC ATC AAC TAT GAT GAG	1824
Gln Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu	
595 600 605	
TTT CCC ACC ATG GTG TTT CCT TCT GGG CAG ATC AGC CAG GCC TCG GCC	1872
Phe Pro Thr Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala	
610 615 620	
TTG GCC CCG GCC CCT CCC CAA GTC CTG CCC CAG GCT CCA GCC CCT GCC	1920
Leu Ala Pro Ala Pro Pro Gln Val Leu Pro Gln Ala Pro Ala Pro Ala	
625 630 635 640	
CCT GCT CCA GCC ATG GTA TCA GCT CTG GCC CAG GCC CCA GCC CCT GTC	1968
Pro Ala Pro Ala Met Val Ser Ala Leu Ala Gln Ala Pro Ala Pro Val	
645 650 655	
CCA GTC CTA GCC CCA GGC CCT CCT CAG GCT GTG GCC CCA CCT GCC CCC	2016
Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val Ala Pro Pro Ala Pro	
660 665 670	
AAG CCC ACC CAG GCT GGG GAA GGA ACG CTG TCA GAG GCC CTG CTG CAG	2064
Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln	
675 680 685	
CTG CAG TTT GAT GAT GAA GAC CTG GGG GCC TTG CTT GGC AAC AGC ACA	2112
Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr	
690 695 700	
GAC CCA GCT GTG TTC ACA GAC CTG GCA TCC GTC GAC AAC TCC GAG TTT	2160
Asp Pro Ala Val Phe Thr Asp Leu Ala Ser Val Asp Asn Ser Glu Phe	
705 710 715 720	
CAG CAG CTG CTG AAC CAG GGC ATA CCT GTG GCC CCC CAC ACA ACT GAG	2208
Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr Glu	
725 730 735	
CCC ATG CTG ATG GAG TAC CCT GAG GCT ATA ACT CGC CTA GTG ACA GGG	2256
Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr Arg Leu Val Thr Gly	
740 745 750	

GCC CAG AGG CCC CCC GAC CCA GCT CCT GCT CCA CTG GGG GCC CCG GGG 2304
 Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly
 755 760 765

CTC CCC AAT GGC CTC CTT TCA GGA GAT GAA GAC TTC TCC TCC ATT GCG 2352
 Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala
 770 775 780

GAC ATG GAC TTC TCA GCC CTG CTG AGT CAG ATC AGC TCC TAA 2394
 Asp Met Asp Phe Ser Ala Leu Leu Ser Gln Ile Ser Ser
 785 790 795

(2) INFORMATION FOR SEQ ID NO:143:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 797 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 1 5 10 15
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 20 25 30
 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 35 40 45
 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 50 55 60
 Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 65 70 75 80
 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 85 90 95
 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 100 105 110
 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 115 120 125
 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 130 135 140
 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
 145 150 155 160
 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
 165 170 175
 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
 180 185 190
 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
 195 200 205
 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
 210 215 220
 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser
 225 230 235 240
 Gly Leu Arg Ser Arg Ala Met Asp Glu Leu Phe Pro Leu Ile Phe Pro

245										250					255					
Ala	Glu	Pro	Ala	Gln	Ala	Ser	Gly	Pro	Tyr	Val	Glu	Ile	Ile	Glu	Gln					
260										265					270					
Pro	Lys	Gln	Arg	Gly	Met	Arg	Phe	Arg	Tyr	Lys	Cys	Glu	Gly	Arg	Ser					
275										280					285					
Ala	Gly	Ser	Ile	Pro	Gly	Glu	Arg	Ser	Thr	Asp	Thr	Thr	Lys	Thr	His					
290										295					300					
Pro	Thr	Ile	Lys	Ile	Asn	Gly	Tyr	Thr	Gly	Pro	Gly	Thr	Val	Arg	Ile					
305	310										315					320				
Ser	Leu	Val	Thr	Lys	Asp	Pro	Pro	His	Arg	Pro	His	Pro	His	Glu	Leu					
325										330					335					
Val	Gly	Lys	Asp	Cys	Arg	Asp	Gly	Phe	Tyr	Glu	Ala	Glu	Leu	Cys	Pro					
340										345					350					
Asp	Arg	Cys	Ile	His	Ser	Phe	Gln	Asn	Leu	Gly	Ile	Gln	Cys	Val	Lys					
355										360					365					
Lys	Arg	Asp	Leu	Glu	Gln	Ala	Ile	Ser	Gln	Arg	Ile	Gln	Thr	Asn	Asn					
370										375					380					
Asn	Pro	Phe	Gln	Val	Pro	Ile	Glu	Glu	Gln	Arg	Gly	Asp	Tyr	Asp	Leu					
385	390										395					400				
Asn	Ala	Val	Arg	Leu	Cys	Phe	Gln	Val	Thr	Val	Arg	Asp	Pro	Ser	Gly					
405										410					415					
Arg	Pro	Leu	Arg	Leu	Pro	Pro	Val	Leu	Pro	His	Pro	Ile	Phe	Asp	Asn					
420										425					430					
Arg	Ala	Pro	Asn	Thr	Ala	Glu	Leu	Lys	Ile	Cys	Arg	Val	Asn	Arg	Asn					
435										440					445					
Ser	Gly	Ser	Cys	Leu	Gly	Gly	Asp	Glu	Ile	Phe	Leu	Leu	Cys	Asp	Lys					
450										455					460					
Val	Gln	Lys	Glu	Asp	Ile	Glu	Val	Tyr	Phe	Thr	Gly	Pro	Gly	Trp	Glu					
465	470										475					480				
Ala	Arg	Gly	Ser	Phe	Ser	Gln	Ala	Asp	Val	His	Arg	Gln	Val	Ala	Ile					
485										490					495					
Val	Phe	Arg	Thr	Pro	Pro	Tyr	Ala	Asp	Pro	Ser	Leu	Gln	Ala	Pro	Val					
500										505					510					
Arg	Val	Ser	Met	Gln	Leu	Arg	Arg	Pro	Ser	Asp	Arg	Glu	Leu	Ser	Glu					
515										520					525					
Pro	Met	Glu	Phe	Gln	Tyr	Leu	Pro	Asp	Thr	Asp	Asp	Arg	His	Arg	Ile					
530										535					540					
Glu	Glu	Lys	Arg	Lys	Arg	Thr	Tyr	Glu	Thr	Phe	Lys	Ser	Ile	Met	Lys					
545	550										555					560				
Lys	Ser	Pro	Phe	Ser	Gly	Pro	Thr	Asp	Pro	Arg	Pro	Pro	Pro	Arg	Arg					
565										570					575					
Ile	Ala	Val	Pro	Ser	Arg	Ser	Ser	Ala	Ser	Val	Pro	Lys	Pro	Ala	Pro					
580										585					590					
Gln	Pro	Tyr	Pro	Phe	Thr	Ser	Ser	Leu	Ser	Thr	Ile	Asn	Tyr	Asp	Glu					
595										600					605					
Phe	Pro	Thr	Met	Val	Phe	Pro	Ser	Gly	Gln	Ile	Ser	Gln	Ala	Ser	Ala					
610										615					620					
Leu	Ala	Pro	Ala	Pro	Pro	Gln	Val	Leu	Pro	Gln	Ala	Pro	Ala	Pro	Ala					
625	630										635					640				
Pro	Ala	Pro	Ala	Met	Val	Ser	Ala	Leu	Ala	Gln	Ala	Pro	Ala	Pro	Val					
645										650					655					
Pro	Val	Leu	Ala	Pro	Gly	Pro	Pro	Gln	Ala	Val	Ala	Pro	Pro	Ala	Pro					
660										665					670					
Lys	Pro	Thr	Gln	Ala	Gly	Glu	Gly	Thr	Leu	Ser	Glu	Ala	Leu	Leu	Gln					
675																				

705		710		715		720
Gln	Gln	Leu	Leu	Asn	Gln	Gly
		725		730		735
Pro	Met	Leu	Met	Glu	Tyr	Pro
		740		745		750
Ala	Gln	Arg	Pro	Pro	Asp	Pro
		755		760		765
Leu	Pro	Asn	Gly	Leu	Leu	Ser
		770		775		780
Asp	Met	Asp	Phe	Ser	Ala	Leu
		785		790		795

(2) INFORMATION FOR SEQ ID NO:144:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3381 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...3378
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

ATG	GAG	CGG	GCC	GGC	CCC	AGC	TTC	GGG	CAG	CAG	CGA	CAG	CAG	CAG		48
Met	Glu	Arg	Ala	Gly	Pro	Ser	Phe	Gly	Gln	Gln	Arg	Gln	Gln	Gln	Gln	
1				5				10						15		
CCC	CAG	CAG	CAG	AAG	CAG	CAG	CAG	AGG	GAT	CAG	GAC	TCG	GTC	GAA	GCA	96
Pro	Gln	Gln	Gln	Lys	Gln	Gln	Gln	Arg	Asp	Gln	Asp	Ser	Val	Glu	Ala	
			20					25						30		
TGG	CTG	GAC	GAT	CAC	TGG	GAC	TTT	ACC	TTC	TCA	TAC	TTT	GTT	AGA	AAA	144
Trp	Leu	Asp	Asp	His	Trp	Asp	Phe	Thr	Phe	Ser	Tyr	Phe	Val	Arg	Lys	
		35					40					45				
GCC	ACC	AGA	GAA	ATG	GTC	AAT	GCA	TGG	TTT	GCT	GAG	AGA	GTT	CAC	ACC	192
Ala	Thr	Arg	Glu	Met	Val	Asn	Ala	Trp	Phe	Ala	Glu	Arg	Val	His	Thr	
		50				55				60						
ATC	CCT	GTG	TGC	AAG	GAA	GGT	ATC	AGA	GGC	CAC	ACC	GAA	TCT	TGC	TCT	240
Ile	Pro	Val	Cys	Lys	Glu	Gly	Ile	Arg	Gly	His	Thr	Glu	Ser	Cys	Ser	
65				70				75				80				
TGT	CCC	TTG	CAG	CAG	AGT	CCT	CGT	GCA	GAT	AAC	AGT	GTC	CCT	GGA	ACA	288
Cys	Pro	Leu	Gln	Gln	Ser	Pro	Arg	Ala	Asp	Asn	Ser	Val	Pro	Gly	Thr	

85	90	95	
CCA ACC AGG AAA ATC TCT GCC TCT GAA TTT GAC CGG CCT CTT AGA CCC Pro Thr Arg Lys Ile Ser Ala Ser Glu Phe Asp Arg Pro Leu Arg Pro 100 105 110			336
ATT GTT GTC AAG GAT TCT GAG GGA ACT GTG AGC TTC CTC TCT GAC TCA Ile Val Val Lys Asp Ser Glu Gly Thr Val Ser Phe Leu Ser Asp Ser 115 120 125			384
GAA AAG AAG GAA CAG ATG CCT CTA ACC CCT CCA AGG TTT GAT CAT GAT Glu Lys Lys Glu Gln Met Pro Leu Thr Pro Pro Arg Phe Asp His Asp 130 135 140			432
GAA GGG GAC CAG TGC TCA AGA CTC TTG GAA TTA GTG AAG GAT ATT TCT Glu Gly Asp Gln Cys Ser Arg Leu Leu Glu Leu Val Lys Asp Ile Ser 145 150 155 160			480
AGT CAT TTG GAT GTC ACA GCC TTA TGT CAC AAA ATT TTC TTG CAT ATC Ser His Leu Asp Val Thr Ala Leu Cys His Lys Ile Phe Leu His Ile 165 170 175			528
CAT GGA CTG ATA TCT GCT GAC CGC TAT TCC CTG TTC CTT GTC TGT GAA His Gly Leu Ile Ser Ala Asp Arg Tyr Ser Leu Phe Leu Val Cys Glu 180 185 190			576
GAC AGC TCC AAT GAC AAG TTT CTT ATC AGC CGC CTC TTT GAT GTT GCT Asp Ser Ser Asn Asp Lys Phe Leu Ile Ser Arg Leu Phe Asp Val Ala 195 200 205			624
GAA GGT TCA ACA CTG GAA GAA GTT TCA AAT AAC TGT ATC CGC TTA GAA Glu Gly Ser Thr Leu Glu Glu Val Ser Asn Asn Cys Ile Arg Leu Glu 210 215 220			672
TGG AAC AAA GGC ATT GTG GGA CAT GTG GCA GCG CTT GGT GAG CCC TTG Trp Asn Lys Gly Ile Val Gly His Val Ala Ala Leu Gly Glu Pro Leu 225 230 235 240			720
AAC ATC AAA GAT GCA TAT GAG GAT CCT CGG TTC AAT GCA GAA GTT GAC Asn Ile Lys Asp Ala Tyr Glu Asp Pro Arg Phe Asn Ala Glu Val Asp 245 250 255			768
CAA ATT ACA GGC TAC AAG ACA CAA AGC ATT CTT TGT ATG CCA ATT AAG Gln Ile Thr Gly Tyr Lys Thr Gln Ser Ile Leu Cys Met Pro Ile Lys 260 265 270			816
AAT CAT AGG GAA GAG GTT GTT GGT GTA GCC CAG GCC ATC AAC AAG AAA Asn His Arg Glu Glu Val Val Gly Val Ala Gln Ala Ile Asn Lys Lys 275 280 285			864
TCA GGA AAC GGT GGG ACA TTT ACT GAA AAA GAT GAA AAG GAC TTT GCT Ser Gly Asn Gly Gly Thr Phe Thr Glu Lys Asp Glu Lys Asp Phe Ala 290 295 300			912

GCT	TAT	TTG	GCA	TTT	TGT	GGT	ATT	GTT	CTT	CAT	AAT	GCT	CAG	CTC	TAT	960
Ala	Tyr	Leu	Ala	Phe	Cys	Gly	Ile	Val	Leu	His	Asn	Ala	Gln	Leu	Tyr	
305					310					315					320	
GAG	ACT	TCA	CTG	CTG	GAG	AAC	AAG	AGA	AAT	CAG	GTG	CTG	CTT	GAC	CTT	1008
Glu	Thr	Ser	Leu	Leu	Glu	Asn	Lys	Arg	Asn	Gln	Val	Leu	Leu	Asp	Leu	
				325					330					335		
GCT	AGT	TTA	ATT	TTT	GAA	GAA	CAA	CAA	TCA	TTA	GAA	GTA	ATT	TTG	AAG	1056
Ala	Ser	Leu	Ile	Phe	Glu	Glu	Gln	Gln	Ser	Leu	Glu	Val	Ile	Leu	Lys	
			340					345					350			
AAA	ATA	GCT	GCC	ACT	ATT	ATC	TCT	TTC	ATG	CAA	GTG	CAG	AAA	TGC	ACC	1104
Lys	Ile	Ala	Ala	Thr	Ile	Ile	Ser	Phe	Met	Gln	Val	Gln	Lys	Cys	Thr	
		355					360				365					
ATT	TTC	ATA	GTG	GAT	GAA	GAT	TGC	TCC	GAT	TCT	TTT	TCT	AGT	GTG	TTT	1152
Ile	Phe	Ile	Val	Asp	Glu	Asp	Cys	Ser	Asp	Ser	Phe	Ser	Ser	Val	Phe	
	370					375				380						
CAC	ATG	GAG	TGT	GAG	GAA	TTA	GAA	AAA	TCA	TCT	GAT	ACA	TTA	ACA	AGG	1200
His	Met	Glu	Cys	Glu	Glu	Leu	Glu	Lys	Ser	Ser	Asp	Thr	Leu	Thr	Arg	
385					390					395					400	
GAA	CAT	GAT	GCA	AAC	AAA	ATC	AAT	TAC	ATG	TAT	GCT	CAG	TAT	GTC	AAA	1248
Glu	His	Asp	Ala	Asn	Lys	Ile	Asn	Tyr	Met	Tyr	Ala	Gln	Tyr	Val	Lys	
			405					410					415			
AAT	ACT	ATG	GAA	CCA	CTT	AAT	ATC	CCA	GAT	GTC	AGT	AAG	GAT	AAA	AGA	1296
Asn	Thr	Met	Glu	Pro	Leu	Asn	Ile	Pro	Asp	Val	Ser	Lys	Asp	Lys	Arg	
		420						425				430				
TTT	CCC	TGG	ACA	ACT	GAA	AAT	ACA	GGA	AAT	GTA	AAC	CAG	CAG	TGC	ATT	1344
Phe	Pro	Trp	Thr	Thr	Glu	Asn	Thr	Gly	Asn	Val	Asn	Gln	Gln	Cys	Ile	
	435						440				445					
AGA	AGT	TTG	CTT	TGT	ACA	CCT	ATA	AAA	AAT	GGA	AAG	AAG	AAT	AAA	GTT	1392
Arg	Ser	Leu	Leu	Cys	Thr	Pro	Ile	Lys	Asn	Gly	Lys	Lys	Asn	Lys	Val	
	450					455				460						
ATA	GGG	GTT	TGC	CAA	CTT	GTT	AAT	AAG	ATG	GAG	GAG	AAT	ACT	GGC	AAG	1440
Ile	Gly	Val	Cys	Gln	Leu	Val	Asn	Lys	Met	Glu	Glu	Asn	Thr	Gly	Lys	
465					470				475					480		
GTT	AAG	CCT	TTC	AAC	CGA	AAT	GAC	GAA	CAG	TTT	CTG	GAA	GCT	TTT	GTC	1488
Val	Lys	Pro	Phe	Asn	Arg	Asn	Asp	Glu	Gln	Phe	Leu	Glu	Ala	Phe	Val	
			485					490					495			
ATC	TTT	TGT	GGC	TTG	GGG	ATC	CAG	AAC	ACG	CAG	ATG	TAT	GAA	GCA	GTG	1536
Ile	Phe	Cys	Gly	Leu	Gly	Ile	Gln	Asn	Thr	Gln	Met	Tyr	Glu	Ala	Val	
		500					505				510					
GAG	AGA	GCC	ATG	GCC	AAG	CAA	ATG	GTC	ACA	TTG	GAG	GTT	CTG	TCG	TAT	1584

515	520	525	
CAT GCT TCA GCA GCA GAG GAA GAA ACA AGA GAG CTA CAG TCG TTA GCG			1632
His Ala Ser Ala Ala Glu Glu Glu Thr Arg Glu Leu Gln Ser Leu Ala			
530	535	540	
GCT GCT GTG GTG CCA TCT GCC CAG ACC CTT AAA ATT ACT GAC TTT AGC			1680
Ala Ala Val Val Pro Ser Ala Gln Thr Leu Lys Ile Thr Asp Phe Ser			
545	550	555	560
TTC AGT GAC TTT GAG CTG TCT GAT CTG GAA ACA GCA CTG TGC ACA ATT			1728
Phe Ser Asp Phe Glu Leu Ser Asp Leu Glu Thr Ala Leu Cys Thr Ile			
565	570	575	
CGG ATG TTT ACT GAC CTC AAC CTT GTG CAG AAC TTC CAG ATG AAA CAT			1776
Arg Met Phe Thr Asp Leu Asn Leu Val Gln Asn Phe Gln Met Lys His			
580	585	590	
GAG GTT CTT TGC AGA TGG ATT TTA AGT GTT AAG AAG AAT TAT CGG AAG			1824
Glu Val Leu Cys Arg Trp Ile Leu Ser Val Lys Lys Asn Tyr Arg Lys			
595	600	605	
AAT GTT GCC TAT CAT AAT TGG AGA CAT GCC TTT AAT ACA GCT CAG TGC			1872
Asn Val Ala Tyr His Asn Trp Arg His Ala Phe Asn Thr Ala Gln Cys			
610	615	620	
ATG TTT GCT GCT CTA AAA GCA GGC AAA ATT CAG AAC AAG CTG ACT GAC			1920
Met Phe Ala Ala Leu Lys Ala Gly Lys Ile Gln Asn Lys Leu Thr Asp			
625	630	635	640
CTG GAG ATA CTT GCA TTG CTG ATT GCT GCA CTA AGC CAC GAT TTG GAT			1968
Leu Glu Ile Leu Ala Leu Leu Ile Ala Ala Leu Ser His Asp Leu Asp			
645	650	655	
CAC CGT GGT GTG AAT AAC TCT TAC ATA CAG CGA AGT GAA CAT CCA CTT			2016
His Arg Gly Val Asn Asn Ser Tyr Ile Gln Arg Ser Glu His Pro Leu			
660	665	670	
GCC CAG CTT TAC TGC CAT TCA ATC ATG GAA CAC CAT CAT TTT GAC CAG			2064
Ala Gln Leu Tyr Cys His Ser Ile Met Glu His His His Phe Asp Gln			
675	680	685	
TGC CTG ATG ATT CTT AAT AGT CCA GGC AAT CAG ATT CTC AGT GGC CTC			2112
Cys Leu Met Ile Leu Asn Ser Pro Gly Asn Gln Ile Leu Ser Gly Leu			
690	695	700	
TCC ATT GAA GAA TAT AAG ACC ACG TTG AAA ATA ATC AAG CAA GCT ATT			2160
Ser Ile Glu Glu Tyr Lys Thr Thr Leu Lys Ile Ile Lys Gln Ala Ile			
705	710	715	720
TTA GCT ACA GAC CTA GCA CTG TAC ATT AAG AGG CGA GGA GAA TTT TTT			2208
Leu Ala Thr Asp Leu Ala Leu Tyr Ile Lys Arg Arg Gly Glu Phe Phe			
725	730	735	

GAA CTT ATA AGA AAA AAT CAA TTC AAT TTG GAA GAT CCT CAT CAA AAG Glu Leu Ile Arg Lys Asn Gln Phe Asn Leu Glu Asp Pro His Gln Lys 740 745 750	2256
GAG TTG TTT TTG GCA ATG CTG ATG ACA GCT TGT GAT CTT TCT GCA ATT Glu Leu Phe Leu Ala Met Leu Met Thr Ala Cys Asp Leu Ser Ala Ile 755 760 765	2304
ACA AAA CCC TGG CCT ATT CAA CAA CGG ATA GCA GAA CTT GTA GCA ACT Thr Lys Pro Trp Pro Ile Gln Gln Arg Ile Ala Glu Leu Val Ala Thr 770 775 780	2352
GAA TTT TTT GAT CAA GGA GAC AGA GAG AGA AAA GAA CTC AAC ATA GAA Glu Phe Phe Asp Gln Gly Asp Arg Glu Arg Lys Glu Leu Asn Ile Glu 785 790 795 800	2400
CCC ACT GAT CTA ATG AAC AGG GAG AAG AAA AAC AAA ATC CCA AGT ATG Pro Thr Asp Leu Met Asn Arg Glu Lys Lys Asn Lys Ile Pro Ser Met 805 810 815	2448
CAA GTT GGG TTC ATA GAT GCC ATC TGC TTG CAA CTG TAT GAG GCC CTG Gln Val Gly Phe Ile Asp Ala Ile Cys Leu Gln Leu Tyr Glu Ala Leu 820 825 830	2496
ACC CAC GTG TCA GAG GAC TGT TTC CCT TTG CTA GAT GGC TGC AGA AAG Thr His Val Ser Glu Asp Cys Phe Pro Leu Leu Asp Gly Cys Arg Lys 835 840 845	2544
AAC AGG CAG AAA TGG CAG GCC CTT GCA GAA CAG CAG GAG AAG ATG CTG Asn Arg Gln Lys Trp Gln Ala Leu Ala Glu Gln Gln Glu Lys Met Leu 850 855 860	2592
ATT AAT GGG GAA AGC GGC CAG GCC AAG CGG AAC TGG GTA CCG CGG GCC Ile Asn Gly Glu Ser Gly Gln Ala Lys Arg Asn Trp Val Pro Arg Ala 865 870 875 880	2640
CGG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC Arg Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe 885 890 895	2688
ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly 900 905 910	2736
CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly 915 920 925	2784
AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro 930 935 940	2832
TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser	2880

945	950	955	960	
CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG				2928
Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met				
965		970	975	
CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC				2976
Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly				
980	985		990	
AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG				3024
Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val				
995	1000		1005	
AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC				3072
Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile				
1010	1015		1020	
CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC				3120
Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile				
1025	1030	1035	1040	
ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC				3168
Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg				
1045	1050		1055	
CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG				3216
His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln				
1060	1065		1070	
AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC				3264
Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr				
1075	1080		1085	
CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT				3312
Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp				
1090	1095		1100	
CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC				3360
His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly				
1105	1110	1115	1120	
ATG GAC GAG CTG TAC AAG TAA				3381
Met Asp Glu Leu Tyr Lys				
1125				

(2) INFORMATION FOR SEQ ID NO:145:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1126 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

```

Met Glu Arg Ala Gly Pro Ser Phe Gly Gln Gln Arg Gln Gln Gln Gln
 1           5           10           15
Pro Gln Gln Gln Lys Gln Gln Gln Arg Asp Gln Asp Ser Val Glu Ala
      20           25           30
Trp Leu Asp Asp His Trp Asp Phe Thr Phe Ser Tyr Phe Val Arg Lys
      35           40           45
Ala Thr Arg Glu Met Val Asn Ala Trp Phe Ala Glu Arg Val His Thr
      50           55           60
Ile Pro Val Cys Lys Glu Gly Ile Arg Gly His Thr Glu Ser Cys Ser
      65           70           75           80
Cys Pro Leu Gln Gln Ser Pro Arg Ala Asp Asn Ser Val Pro Gly Thr
      85           90           95
Pro Thr Arg Lys Ile Ser Ala Ser Glu Phe Asp Arg Pro Leu Arg Pro
      100          105          110
Ile Val Val Lys Asp Ser Glu Gly Thr Val Ser Phe Leu Ser Asp Ser
      115          120          125
Glu Lys Lys Glu Gln Met Pro Leu Thr Pro Pro Arg Phe Asp His Asp
      130          135          140
Glu Gly Asp Gln Cys Ser Arg Leu Leu Glu Leu Val Lys Asp Ile Ser
      145          150          155          160
Ser His Leu Asp Val Thr Ala Leu Cys His Lys Ile Phe Leu His Ile
      165          170          175
His Gly Leu Ile Ser Ala Asp Arg Tyr Ser Leu Phe Leu Val Cys Glu
      180          185          190
Asp Ser Ser Asn Asp Lys Phe Leu Ile Ser Arg Leu Phe Asp Val Ala
      195          200          205
Glu Gly Ser Thr Leu Glu Glu Val Ser Asn Asn Cys Ile Arg Leu Glu
      210          215          220
Trp Asn Lys Gly Ile Val Gly His Val Ala Ala Leu Gly Glu Pro Leu
      225          230          235          240
Asn Ile Lys Asp Ala Tyr Glu Asp Pro Arg Phe Asn Ala Glu Val Asp
      245          250          255
Gln Ile Thr Gly Tyr Lys Thr Gln Ser Ile Leu Cys Met Pro Ile Lys
      260          265          270
Asn His Arg Glu Glu Val Val Gly Val Ala Gln Ala Ile Asn Lys Lys
      275          280          285
Ser Gly Asn Gly Gly Thr Phe Thr Glu Lys Asp Glu Lys Asp Phe Ala
      290          295          300
Ala Tyr Leu Ala Phe Cys Gly Ile Val Leu His Asn Ala Gln Leu Tyr
      305          310          315          320
Glu Thr Ser Leu Leu Glu Asn Lys Arg Asn Gln Val Leu Leu Asp Leu
      325          330          335
Ala Ser Leu Ile Phe Glu Glu Gln Gln Ser Leu Glu Val Ile Leu Lys
      340          345          350
Lys Ile Ala Ala Thr Ile Ile Ser Phe Met Gln Val Gln Lys Cys Thr
      355          360          365
Ile Phe Ile Val Asp Glu Asp Cys Ser Asp Ser Phe Ser Ser Val Phe
      370          375          380

```

His Met Glu Cys Glu Glu Leu Glu Lys Ser Ser Asp Thr Leu Thr Arg
 385 390 395 400
 Glu His Asp Ala Asn Lys Ile Asn Tyr Met Tyr Ala Gln Tyr Val Lys
 405 410 415
 Asn Thr Met Glu Pro Leu Asn Ile Pro Asp Val Ser Lys Asp Lys Arg
 420 425 430
 Phe Pro Trp Thr Thr Glu Asn Thr Gly Asn Val Asn Gln Gln Cys Ile
 435 440 445
 Arg Ser Leu Leu Cys Thr Pro Ile Lys Asn Gly Lys Lys Asn Lys Val
 450 455 460
 Ile Gly Val Cys Gln Leu Val Asn Lys Met Glu Glu Asn Thr Gly Lys
 465 470 475 480
 Val Lys Pro Phe Asn Arg Asn Asp Glu Gln Phe Leu Glu Ala Phe Val
 485 490 495
 Ile Phe Cys Gly Leu Gly Ile Gln Asn Thr Gln Met Tyr Glu Ala Val
 500 505 510
 Glu Arg Ala Met Ala Lys Gln Met Val Thr Leu Glu Val Leu Ser Tyr
 515 520 525
 His Ala Ser Ala Ala Glu Glu Glu Thr Arg Glu Leu Gln Ser Leu Ala
 530 535 540
 Ala Ala Val Val Pro Ser Ala Gln Thr Leu Lys Ile Thr Asp Phe Ser
 545 550 555 560
 Phe Ser Asp Phe Glu Leu Ser Asp Leu Glu Thr Ala Leu Cys Thr Ile
 565 570 575
 Arg Met Phe Thr Asp Leu Asn Leu Val Gln Asn Phe Gln Met Lys His
 580 585 590
 Glu Val Leu Cys Arg Trp Ile Leu Ser Val Lys Lys Asn Tyr Arg Lys
 595 600 605
 Asn Val Ala Tyr His Asn Trp Arg His Ala Phe Asn Thr Ala Gln Cys
 610 615 620
 Met Phe Ala Ala Leu Lys Ala Gly Lys Ile Gln Asn Lys Leu Thr Asp
 625 630 635 640
 Leu Glu Ile Leu Ala Leu Leu Ile Ala Ala Leu Ser His Asp Leu Asp
 645 650 655
 His Arg Gly Val Asn Asn Ser Tyr Ile Gln Arg Ser Glu His Pro Leu
 660 665 670
 Ala Gln Leu Tyr Cys His Ser Ile Met Glu His His His Phe Asp Gln
 675 680 685
 Cys Leu Met Ile Leu Asn Ser Pro Gly Asn Gln Ile Leu Ser Gly Leu
 690 695 700
 Ser Ile Glu Glu Tyr Lys Thr Thr Leu Lys Ile Ile Lys Gln Ala Ile
 705 710 715 720
 Leu Ala Thr Asp Leu Ala Leu Tyr Ile Lys Arg Arg Gly Glu Phe Phe
 725 730 735
 Glu Leu Ile Arg Lys Asn Gln Phe Asn Leu Glu Asp Pro His Gln Lys
 740 745 750
 Glu Leu Phe Leu Ala Met Leu Met Thr Ala Cys Asp Leu Ser Ala Ile
 755 760 765
 Thr Lys Pro Trp Pro Ile Gln Gln Arg Ile Ala Glu Leu Val Ala Thr
 770 775 780
 Glu Phe Phe Asp Gln Gly Asp Arg Glu Arg Lys Glu Leu Asn Ile Glu
 785 790 795 800
 Pro Thr Asp Leu Met Asn Arg Glu Lys Lys Asn Lys Ile Pro Ser Met
 805 810 815

Gln Val Gly Phe Ile Asp Ala Ile Cys Leu Gln Leu Tyr Glu Ala Leu
 820 825 830
 Thr His Val Ser Glu Asp Cys Phe Pro Leu Leu Asp Gly Cys Arg Lys
 835 840 845
 Asn Arg Gln Lys Trp Gln Ala Leu Ala Glu Gln Gln Glu Lys Met Leu
 850 855 860
 Ile Asn Gly Glu Ser Gly Gln Ala Lys Arg Asn Trp Val Pro Arg Ala
 865 870 875 880
 Arg Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe
 885 890 895
 Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly
 900 905 910
 His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly
 915 920 925
 Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro
 930 935 940
 Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser
 945 950 955 960
 Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met
 965 970 975
 Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly
 980 985 990
 Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val
 995 1000 1005
 Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile
 1010 1015 1020
 Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile
 1025 1030 1035 1040
 Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg
 1045 1050 1055
 His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln
 1060 1065 1070
 Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr
 1075 1080 1085
 Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp
 1090 1095 1100
 His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly
 1105 1110 1115 1120
 Met Asp Glu Leu Tyr Lys
 1125

(2) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2760 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2757

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

ATG GCT GAC CCG GCT GCG GGG CCG CCG CCG AGC GAG GGC GAG GAG AGC	48
Met Ala Asp Pro Ala Ala Gly Pro Pro Pro Ser Glu Gly Glu Glu Ser	
1 5 10 15	
ACC GTG CGC TTC GCC CGC AAA GGC GCC CTC CGG CAG AAG AAC GTG CAT	96
Thr Val Arg Phe Ala Arg Lys Gly Ala Leu Arg Gln Lys Asn Val His	
20 25 30	
GAG GTC AAG AAC CAC AAA TTC ACC GCC CGC TTC TTC AAG CAG CCC ACC	144
Glu Val Lys Asn His Lys Phe Thr Ala Arg Phe Phe Lys Gln Pro Thr	
35 40 45	
TTC TGC AGC CAC TGC ACC GAC TTC ATC TGG GGC TTC GGG AAG CAG GGA	192
Phe Cys Ser His Cys Thr Asp Phe Ile Trp Gly Phe Gly Lys Gln Gly	
50 55 60	
TTC CAG TGC CAA GTT TGC TGC TTT GTG GTG CAC AAG CGG TGC CAT GAA	240
Phe Gln Cys Gln Val Cys Cys Phe Val Val His Lys Arg Cys His Glu	
65 70 75 80	
TTT GTC ACA TTC TCC TGC CCT GGC GCT GAC AAG GGT CCA GCC TCC GAT	288
Phe Val Thr Phe Ser Cys Pro Gly Ala Asp Lys Gly Pro Ala Ser Asp	
85 90 95	
GAC CCC CGC AGC AAA CAC AAG TTT AAG ATC CAC ACG TAC TCC AGC CCC	336
Asp Pro Arg Ser Lys His Lys Phe Lys Ile His Thr Tyr Ser Ser Pro	
100 105 110	
ACG TTT TGT GAC CAC TGT GGG TCA CTG CTG TAT GGA CTC ATC CAC CAG	384
Thr Phe Cys Asp His Cys Gly Ser Leu Leu Tyr Gly Leu Ile His Gln	
115 120 125	
GGG ATG AAA TGT GAC ACC TGC ATG ATG AAT GTG CAC AAG CGC TGC GTG	432
Gly Met Lys Cys Asp Thr Cys Met Met Asn Val His Lys Arg Cys Val	
130 135 140	
ATG AAT GTT CCC AGC CTG TGT GGC ACG GAC CAC ACG GAG CGC CGC GGC	480
Met Asn Val Pro Ser Leu Cys Gly Thr Asp His Thr Glu Arg Arg Gly	
145 150 155 160	
CGC ATC TAC ATC CAG GCC CAC ATC GAC AGG GAC GTC CTC ATT GTC CTC	528
Arg Ile Tyr Ile Gln Ala His Ile Asp Arg Asp Val Leu Ile Val Leu	
165 170 175	
GTA AGA GAT GCT AAA AAC CTT GTA CCT ATG GAC CCC AAT GGC CTG TCA	576
Val Arg Asp Ala Lys Asn Leu Val Pro Met Asp Pro Asn Gly Leu Ser	
180 185 190	
GAT CCC TAC GTA AAA CTG AAA CTG ATT CCC GAT CCC AAA AGT GAG AGC	624
Asp Pro Tyr Val Lys Leu Lys Leu Ile Pro Asp Pro Lys Ser Glu Ser	

195	200	205	
AAA CAG AAG ACC AAA ACC ATC AAA TGC TCC CTC AAC CCT GAG TGG AAT Lys Gln Lys Thr Lys Thr Ile Lys Cys Ser Leu Asn Pro Glu Trp Asn 210 215 220			672
GAG ACA TTT AGA TTT CAG CTG AAA GAA TCG GAC AAA GAC AGA AGA CTG Glu Thr Phe Arg Phe Gln Leu Lys Glu Ser Asp Lys Asp Arg Arg Leu 225 230 235 240			720
TCA GTA GAG ATT TGG GAT TGG GAT TTG ACC AGC AGG AAT GAC TTC ATG Ser Val Glu Ile Trp Asp Trp Asp Leu Thr Ser Arg Asn Asp Phe Met 245 250 255			768
GGA TCT TTG TCC TTT GGG ATT TCT GAA CTT CAG AAG GCC AGT GTT GAT Gly Ser Leu Ser Phe Gly Ile Ser Glu Leu Gln Lys Ala Ser Val Asp 260 265 270			816
GGC TGG TTT AAG TTA CTG AGC CAG GAG GAA GGC GAG TAC TTC AAT GTG Gly Trp Phe Lys Leu Leu Ser Gln Glu Glu Gly Glu Tyr Phe Asn Val 275 280 285			864
CCT GTG CCA CCA GAA GGA AGT GAG GCC AAT GAA GAA CTG CGG CAG AAA Pro Val Pro Pro Glu Gly Ser Glu Ala Asn Glu Glu Leu Arg Gln Lys 290 295 300			912
TTT GAG AGG GCC AAG ATC AGT CAG GGA ACC AAG GTC CCG GAA GAA AAG Phe Glu Arg Ala Lys Ile Ser Gln Gly Thr Lys Val Pro Glu Glu Lys 305 310 315 320			960
ACG ACC AAC ACT GTC TCC AAA TTT GAC AAC AAT GGC AAC AGA GAC CGG Thr Thr Asn Thr Val Ser Lys Phe Asp Asn Asn Gly Asn Arg Asp Arg 325 330 335			1008
ATG AAA CTG ACC GAT TTT AAC TTC CTA ATG GTG CTG GGG AAA GGC AGC Met Lys Leu Thr Asp Phe Asn Phe Leu Met Val Leu Gly Lys Gly Ser 340 345 350			1056
TTT GGC AAG GTC ATG CTT TCA GAA CGA AAA GGC ACA GAT GAG CTC TAT Phe Gly Lys Val Met Leu Ser Glu Arg Lys Gly Thr Asp Glu Leu Tyr 355 360 365			1104
GCT GTG AAG ATC CTG AAG AAG GAC GTT GTG ATC CAA GAT GAT GAC GTG Ala Val Lys Ile Leu Lys Lys Asp Val Val Ile Gln Asp Asp Asp Val 370 375 380			1152
GAG TGC ACT ATG GTG GAG AAG CGG GTG TTG GCC CTG CCT GGG AAG CCG Glu Cys Thr Met Val Glu Lys Arg Val Leu Ala Leu Pro Gly Lys Pro 385 390 395 400			1200
CCC TTC CTG ACC CAG CTC CAC TCC TGC TTC CAG ACC ATG GAC CGC CTG Pro Phe Leu Thr Gln Leu His Ser Cys Phe Gln Thr Met Asp Arg Leu 405 410 415			1248

TAC TTT GTG ATG GAG TAC GTG AAT GGG GGC GAC CTC ATG TAT CAC ATC Tyr Phe Val Met Glu Tyr Val Asn Gly Gly Asp Leu Met Tyr His Ile 420 425 430	1296
CAG CAA GTC GGC CGG TTC AAG GAG CCC CAT GCT GTA TTT TAC GCT GCA Gln Gln Val Gly Arg Phe Lys Glu Pro His Ala Val Phe Tyr Ala Ala 435 440 445	1344
GAA ATT GCC ATC GGT CTG TTC TTC TTA CAG AGT AAG GGC ATC ATT TAC Glu Ile Ala Ile Gly Leu Phe Phe Leu Gln Ser Lys Gly Ile Ile Tyr 450 455 460	1392
CGT GAC CTA AAA CTT GAC AAC GTG ATG CTC GAT TCT GAG GGA CAC ATC Arg Asp Leu Lys Leu Asp Asn Val Met Leu Asp Ser Glu Gly His Ile 465 470 475 480	1440
AAG ATT GCC GAT TTT GGC ATG TGT AAG GAA AAC ATC TGG GAT GGG GTG Lys Ile Ala Asp Phe Gly Met Cys Lys Glu Asn Ile Trp Asp Gly Val 485 490 495	1488
ACA ACC AAG ACA TTC TGT GGC ACT CCA GAC TAC ATC GCC CCC GAG ATA Thr Thr Lys Thr Phe Cys Gly Thr Pro Asp Tyr Ile Ala Pro Glu Ile 500 505 510	1536
ATT GCT TAT CAG CCC TAT GGG AAG TCC GTG GAT TGG TGG GCA TTT GGA Ile Ala Tyr Gln Pro Tyr Gly Lys Ser Val Asp Trp Trp Ala Phe Gly 515 520 525	1584
GTC CTG CTG TAT GAA ATG TTG GCT GGG CAG GCA CCC TTT GAA GGG GAG Val Leu Leu Tyr Glu Met Leu Ala Gly Gln Ala Pro Phe Glu Gly Glu 530 535 540	1632
GAT GAA GAT GAA CTC TTC CAA TCC ATC ATG GAA CAC AAC GTA GCC TAT Asp Glu Asp Glu Leu Phe Gln Ser Ile Met Glu His Asn Val Ala Tyr 545 550 555 560	1680
CCC AAG TCT ATG TCC AAG GAA GCT GTG GCC ATC TGC AAA GGG CTG ATG Pro Lys Ser Met Ser Lys Glu Ala Val Ala Ile Cys Lys Gly Leu Met 565 570 575	1728
ACC AAA CAC CCA GGC AAA CGT CTG GGT TGT GGA CCT GAA GGC GAA CGT Thr Lys His Pro Gly Lys Arg Leu Gly Cys Gly Pro Glu Gly Glu Arg 580 585 590	1776
GAT ATC AAA GAG CAT GCA TTT TTC CGG TAT ATT GAT TGG GAG AAA CTT Asp Ile Lys Glu His Ala Phe Phe Arg Tyr Ile Asp Trp Glu Lys Leu 595 600 605	1824
GAA CGC AAA GAG ATC CAG CCC CCT TAT AAG CCA AAA GCT TGT GGG CGA Glu Arg Lys Glu Ile Gln Pro Pro Tyr Lys Pro Lys Ala Cys Gly Arg 610 615 620	1872
AAT GCT GAA AAC TTC GAC CGA TTT TTC ACC CGC CAT CCA CCA GTC CTA Asn Ala Glu Asn Phe Asp Arg Phe Phe Thr Arg His Pro Pro Val Leu	1920

625	630	635	640	
ACA CCT CCC GAC CAG GAA GTC ATC AGG AAT ATT GAC CAA TCA GAA TTC				1968
Thr Pro Pro Asp Gln Glu Val Ile Arg Asn Ile Asp Gln Ser Glu Phe				
645	650	655		
GAA GGA TTT TCC TTT GTT AAC TCT GAA TTT TTA AAA CCC GAA GTC AAG				2016
Glu Gly Phe Ser Phe Val Asn Ser Glu Phe Leu Lys Pro Glu Val Lys				
660	665	670		
AGC TCG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG				2064
Ser Ser Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu				
675	680	685		
TTC ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC				2112
Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn				
690	695	700		
GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC				2160
Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr				
705	710	715	720	
GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG				2208
Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val				
725	730	735		
CCC TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC				2256
Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe				
740	745	750		
AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC				2304
Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala				
755	760	765		
ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC				2352
Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp				
770	775	780		
GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG				2400
Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu				
785	790	795	800	
GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC				2448
Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn				
805	810	815		
ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT				2496
Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr				
820	825	830		
ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC				2544
Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile				
835	840	845		

CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG	2592
Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln	
850 855 860	
CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC	2640
Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His	
865 870 875 880	
TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC	2688
Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg	
885 890 895	
GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC	2736
Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu	
900 905 910	
GGC ATG GAC GAG CTG TAC AAG TAA	2760
Gly Met Asp Glu Leu Tyr Lys	
915	

(2) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 919 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

Met	Ala	Asp	Pro	Ala	Ala	Gly	Pro	Pro	Pro	Ser	Glu	Gly	Glu	Glu	Ser
1				5					10					15	
Thr	Val	Arg	Phe	Ala	Arg	Lys	Gly	Ala	Leu	Arg	Gln	Lys	Asn	Val	His
			20					25					30		
Glu	Val	Lys	Asn	His	Lys	Phe	Thr	Ala	Arg	Phe	Phe	Lys	Gln	Pro	Thr
			35				40					45			
Phe	Cys	Ser	His	Cys	Thr	Asp	Phe	Ile	Trp	Gly	Phe	Gly	Lys	Gln	Gly
			50				55				60				
Phe	Gln	Cys	Gln	Val	Cys	Cys	Phe	Val	Val	His	Lys	Arg	Cys	His	Glu
65				70					75					80	
Phe	Val	Thr	Phe	Ser	Cys	Pro	Gly	Ala	Asp	Lys	Gly	Pro	Ala	Ser	Asp
				85				90					95		
Asp	Pro	Arg	Ser	Lys	His	Lys	Phe	Lys	Ile	His	Thr	Tyr	Ser	Ser	Pro
			100				105						110		
Thr	Phe	Cys	Asp	His	Cys	Gly	Ser	Leu	Leu	Tyr	Gly	Leu	Ile	His	Gln
			115				120					125			
Gly	Met	Lys	Cys	Asp	Thr	Cys	Met	Met	Asn	Val	His	Lys	Arg	Cys	Val
			130				135				140				
Met	Asn	Val	Pro	Ser	Leu	Cys	Gly	Thr	Asp	His	Thr	Glu	Arg	Arg	Gly
145					150				155					160	

Arg Ile Tyr Ile Gln Ala His Ile Asp Arg Asp Val Leu Ile Val Leu
 165 170 175
 Val Arg Asp Ala Lys Asn Leu Val Pro Met Asp Pro Asn Gly Leu Ser
 180 185 190
 Asp Pro Tyr Val Lys Leu Lys Leu Ile Pro Asp Pro Lys Ser Glu Ser
 195 200 205
 Lys Gln Lys Thr Lys Thr Ile Lys Cys Ser Leu Asn Pro Glu Trp Asn
 210 215 220
 Glu Thr Phe Arg Phe Gln Leu Lys Glu Ser Asp Lys Asp Arg Arg Leu
 225 230 235 240
 Ser Val Glu Ile Trp Asp Trp Asp Leu Thr Ser Arg Asn Asp Phe Met
 245 250 255
 Gly Ser Leu Ser Phe Gly Ile Ser Glu Leu Gln Lys Ala Ser Val Asp
 260 265 270
 Gly Trp Phe Lys Leu Leu Ser Gln Glu Glu Gly Glu Tyr Phe Asn Val
 275 280 285
 Pro Val Pro Pro Glu Gly Ser Glu Ala Asn Glu Glu Leu Arg Gln Lys
 290 295 300
 Phe Glu Arg Ala Lys Ile Ser Gln Gly Thr Lys Val Pro Glu Glu Lys
 305 310 315 320
 Thr Thr Asn Thr Val Ser Lys Phe Asp Asn Asn Gly Asn Arg Asp Arg
 325 330 335
 Met Lys Leu Thr Asp Phe Asn Phe Leu Met Val Leu Gly Lys Gly Ser
 340 345 350
 Phe Gly Lys Val Met Leu Ser Glu Arg Lys Gly Thr Asp Glu Leu Tyr
 355 360 365
 Ala Val Lys Ile Leu Lys Lys Asp Val Val Ile Gln Asp Asp Asp Val
 370 375 380
 Glu Cys Thr Met Val Glu Lys Arg Val Leu Ala Leu Pro Gly Lys Pro
 385 390 395 400
 Pro Phe Leu Thr Gln Leu His Ser Cys Phe Gln Thr Met Asp Arg Leu
 405 410 415
 Tyr Phe Val Met Glu Tyr Val Asn Gly Gly Asp Leu Met Tyr His Ile
 420 425 430
 Gln Gln Val Gly Arg Phe Lys Glu Pro His Ala Val Phe Tyr Ala Ala
 435 440 445
 Glu Ile Ala Ile Gly Leu Phe Phe Leu Gln Ser Lys Gly Ile Ile Tyr
 450 455 460
 Arg Asp Leu Lys Leu Asp Asn Val Met Leu Asp Ser Glu Gly His Ile
 465 470 475 480
 Lys Ile Ala Asp Phe Gly Met Cys Lys Glu Asn Ile Trp Asp Gly Val
 485 490 495
 Thr Thr Lys Thr Phe Cys Gly Thr Pro Asp Tyr Ile Ala Pro Glu Ile
 500 505 510
 Ile Ala Tyr Gln Pro Tyr Gly Lys Ser Val Asp Trp Trp Ala Phe Gly
 515 520 525
 Val Leu Leu Tyr Glu Met Leu Ala Gly Gln Ala Pro Phe Glu Gly Glu
 530 535 540
 Asp Glu Asp Glu Leu Phe Gln Ser Ile Met Glu His Asn Val Ala Tyr
 545 550 555 560
 Pro Lys Ser Met Ser Lys Glu Ala Val Ala Ile Cys Lys Gly Leu Met
 565 570 575
 Thr Lys His Pro Gly Lys Arg Leu Gly Cys Gly Pro Glu Gly Glu Arg
 580 585 590

Asp Ile Lys Glu His Ala Phe Phe Arg Tyr Ile Asp Trp Glu Lys Leu
 595 600 605
 Glu Arg Lys Glu Ile Gln Pro Pro Tyr Lys Pro Lys Ala Cys Gly Arg
 610 615 620
 Asn Ala Glu Asn Phe Asp Arg Phe Phe Thr Arg His Pro Pro Val Leu
 625 630 635 640
 Thr Pro Pro Asp Gln Glu Val Ile Arg Asn Ile Asp Gln Ser Glu Phe
 645 650 655
 Glu Gly Phe Ser Phe Val Asn Ser Glu Phe Leu Lys Pro Glu Val Lys
 660 665 670
 Ser Ser Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu
 675 680 685
 Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn
 690 695 700
 Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr
 705 710 715 720
 Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val
 725 730 735
 Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe
 740 745 750
 Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala
 755 760 765
 Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp
 770 775 780
 Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu
 785 790 795 800
 Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn
 805 810 815
 Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr
 820 825 830
 Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile
 835 840 845
 Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln
 850 855 860
 Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His
 865 870 875 880
 Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg
 885 890 895
 Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu
 900 905 910
 Gly Met Asp Glu Leu Tyr Lys
 915

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3009 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
 (B) LOCATION: 1...3006
 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

ATG GCT CAG CAG ACA AGC CCG GAC ACT TTA ACA GTA CCT GAA GTG GAT	48
Met Ala Gln Gln Thr Ser Pro Asp Thr Leu Thr Val Pro Glu Val Asp	
1 5 10 15	
AAT CCG CAT TGT CCA AAC CCG TGG CTG AAC GAA GAC CTT GTG AAA TCC	96
Asn Pro His Cys Pro Asn Pro Trp Leu Asn Glu Asp Leu Val Lys Ser	
20 25 30	
TTG CGA GAA AAC CTG TTG CAG CAT GAG AAG TCC AAG ACA GCG AGG AAA	144
Leu Arg Glu Asn Leu Leu Gln His Glu Lys Ser Lys Thr Ala Arg Lys	
35 40 45	
TCG GTT TCT CCC AAG CTC TCT CCA GTG ATC TCT CCG AGA AAT TCC CCC	192
Ser Val Ser Pro Lys Leu Ser Pro Val Ile Ser Pro Arg Asn Ser Pro	
50 55 60	
AGG CTT CTG CGC AGA ATG CTT CTC AGC AGC AAC ATC CCC AAA CAG CGG	240
Arg Leu Leu Arg Arg Met Leu Leu Ser Ser Asn Ile Pro Lys Gln Arg	
65 70 75 80	
CGT TTC ACG GTG GCA CAT ACA TGT TTT GAT GTG GAC AAT GGC ACA TCT	288
Arg Phe Thr Val Ala His Thr Cys Phe Asp Val Asp Asn Gly Thr Ser	
85 90 95	
GCG GGA CGG AGT CCC TTG GAT CCC ATG ACC AGC CCA GGA TCC GGG CTA	336
Ala Gly Arg Ser Pro Leu Asp Pro Met Thr Ser Pro Gly Ser Gly Leu	
100 105 110	
ATT CTC CAA GCA AAT TTT GTC CAC AGT CAA CGA CGG GAG TCC TTC CTG	384
Ile Leu Gln Ala Asn Phe Val His Ser Gln Arg Arg Glu Ser Phe Leu	
115 120 125	
TAT CGA TCC GAC AGC GAT TAT GAC CTC TCT CCA AAG TCT ATG TCC CGG	432
Tyr Arg Ser Asp Ser Asp Tyr Asp Leu Ser Pro Lys Ser Met Ser Arg	
130 135 140	
AAC TCC TCC ATT GCC AGT GAT ATA CAC GGA GAT GAC TTG ATT GTG ACT	480
Asn Ser Ser Ile Ala Ser Asp Ile His Gly Asp Asp Leu Ile Val Thr	
145 150 155 160	
CCA TTT GCT CAG GTC TTG GCC AGT CTG CGA ACT GTA CGA AAC AAC TTT	528
Pro Phe Ala Gln Val Leu Ala Ser Leu Arg Thr Val Arg Asn Asn Phe	
165 170 175	
GCT GCA TTA ACT AAT TTG CAA GAT CGA GCA CCT AGC AAA AGA TCA CCC	576
Ala Ala Leu Thr Asn Leu Gln Asp Arg Ala Pro Ser Lys Arg Ser Pro	
180 185 190	

ATG TGC AAC CAA CCA TCC ATC AAC AAA GCC ACC ATA ACA GAG GAG GCC Met Cys Asn Gln Pro Ser Ile Asn Lys Ala Thr Ile Thr Glu Glu Ala 195 200 205	624
TAC CAG AAA CTG GCC AGC GAG ACC CTG GAG GAG CTG GAC TGG TGT CTG Tyr Gln Lys Leu Ala Ser Glu Thr Leu Glu Glu Leu Asp Trp Cys Leu 210 215 220	672
GAC CAG CTA GAG ACC CTA CAG ACC AGG CAC TCC GTC AGT GAG ATG GCC Asp Gln Leu Glu Thr Leu Gln Thr Arg His Ser Val Ser Glu Met Ala 225 230 235 240	720
TCC AAC AAG TTT AAA AGG ATG CTT AAT CGG GAG CTC ACC CAT CTC TCT Ser Asn Lys Phe Lys Arg Met Leu Asn Arg Glu Leu Thr His Leu Ser 245 250 255	768
GAA ATG AGT CGG TCT GGA AAT CAA GTG TCA GAG TTT ATA TCA AAC ACA Glu Met Ser Arg Ser Gly Asn Gln Val Ser Glu Phe Ile Ser Asn Thr 260 265 270	816
TTC TTA GAT AAG CAA CAT GAA GTG GAA ATT CCT TCT CCA ACT CAG AAG Phe Leu Asp Lys Gln His Glu Val Glu Ile Pro Ser Pro Thr Gln Lys 275 280 285	864
GAA AAG GAG AAA AAG AAA AGA CCA ATG TCT CAG ATC AGT GGA GTC AAG Glu Lys Glu Lys Lys Lys Arg Pro Met Ser Gln Ile Ser Gly Val Lys 290 295 300	912
AAA TTG ATG CAC AGC TCT AGT CTG ACT AAT TCA AGT ATC CCA AGG TTT Lys Leu Met His Ser Ser Ser Leu Thr Asn Ser Ser Ile Pro Arg Phe 305 310 315 320	960
GGA GTT AAA ACT GAA CAA GAA GAT GTC CTT GCC AAG GAA CTA GAA GAT Gly Val Lys Thr Glu Gln Glu Asp Val Leu Ala Lys Glu Leu Glu Asp 325 330 335	1008
GTG AAC AAA TGG GGT CTT CAT GTT TTC AGA ATA GCA GAG TTG TCT GGT Val Asn Lys Trp Gly Leu His Val Phe Arg Ile Ala Glu Leu Ser Gly 340 345 350	1056
AAC CGG CCC TTG ACT GTT ATC ATG CAC ACC ATT TTT CAG GAA CGG GAT Asn Arg Pro Leu Thr Val Ile Met His Thr Ile Phe Gln Glu Arg Asp 355 360 365	1104
TTA TTA AAA ACA TTT AAA ATT CCA GTA GAT ACT TTA ATT ACA TAT CTT Leu Leu Lys Thr Phe Lys Ile Pro Val Asp Thr Leu Ile Thr Tyr Leu 370 375 380	1152
ATG ACT CTC GAA GAC CAT TAC CAT GCT GAT GTG GCC TAT CAC AAC AAT Met Thr Leu Glu Asp His Tyr His Ala Asp Val Ala Tyr His Asn Asn 385 390 395 400	1200
ATC CAT GCT GCA GAT GTT GTC CAG TCT ACT CAT GTG CTA TTA TCT ACA Ile His Ala Ala Asp Val Val Gln Ser Thr His Val Leu Leu Ser Thr	1248

405	410	415	
CCT GCT TTG GAG GCT GTG TTT ACA GAT TTG GAG ATT CTT GCA GCA ATT			1296
Pro Ala Leu Glu Ala Val Phe Thr Asp Leu Glu Ile Leu Ala Ala Ile			
420	425	430	
TTT GCC AGT GCA ATA CAT GAT GTA GAT CAT CCT GGT GTG TCC AAT CAA			1344
Phe Ala Ser Ala Ile His Asp Val Asp His Pro Gly Val Ser Asn Gln			
435	440	445	
TTT CTG ATC AAT ACA AAC TCT GAA CTT GCC TTG ATG TAC AAT GAT TCC			1392
Phe Leu Ile Asn Thr Asn Ser Glu Leu Ala Leu Met Tyr Asn Asp Ser			
450	455	460	
TCA GTC TTA GAG AAC CAT CAT TTG GCT GTG GGC TTT AAA TTG CTT CAG			1440
Ser Val Leu Glu Asn His His Leu Ala Val Gly Phe Lys Leu Leu Gln			
465	470	475	480
GAA GAA AAC TGT GAC ATT TTC CAG AAT TTG ACC AAA AAA CAA AGA CAA			1488
Glu Glu Asn Cys Asp Ile Phe Gln Asn Leu Thr Lys Lys Gln Arg Gln			
485	490	495	
TCT TTA AGG AAA ATG GTC ATT GAC ATC GTA CTT GCA ACA GAT ATG TCA			1536
Ser Leu Arg Lys Met Val Ile Asp Ile Val Leu Ala Thr Asp Met Ser			
500	505	510	
AAA CAC ATG AAT CTA CTG GCT GAT TTG AAG ACT ATG GTT GAA ACT AAG			1584
Lys His Met Asn Leu Leu Ala Asp Leu Lys Thr Met Val Glu Thr Lys			
515	520	525	
AAA GTG ACA AGC TCT GGA GTT CTT CTT CTT GAT AAT TAT TCC GAT AGG			1632
Lys Val Thr Ser Ser Gly Val Leu Leu Leu Asp Asn Tyr Ser Asp Arg			
530	535	540	
ATT CAG GTT CTT CAG AAT ATG GTG CAC TGT GCA GAT CTG AGC AAC CCA			1680
Ile Gln Val Leu Gln Asn Met Val His Cys Ala Asp Leu Ser Asn Pro			
545	550	555	560
ACA AAG CCT CTC CAG CTG TAC CGC CAG TGG ACG GAC CGG ATA ATG GAG			1728
Thr Lys Pro Leu Gln Leu Tyr Arg Gln Trp Thr Asp Arg Ile Met Glu			
565	570	575	
GAG TTC TTC CGC CAA GGA GAC CGA GAG AGG GAA CGT GGC ATG GAG ATA			1776
Glu Phe Phe Arg Gln Gly Asp Arg Glu Arg Glu Arg Gly Met Glu Ile			
580	585	590	
AGC CCC ATG TGT GAC AAG CAC AAT GCT TCC GTG GAA AAA TCA CAG GTG			1824
Ser Pro Met Cys Asp Lys His Asn Ala Ser Val Glu Lys Ser Gln Val			
595	600	605	
GGC TTC ATA GAC TAT ATT GTT CAT CCC CTC TGG GAG ACA TGG GCA GAC			1872
Gly Phe Ile Asp Tyr Ile Val His Pro Leu Trp Glu Thr Trp Ala Asp			
610	615	620	

CTC GTC CAC CCT GAC GCC CAG GAT ATT TTG GAC ACT TTG GAG GAC AAT Leu Val His Pro Asp Ala Gln Asp Ile Leu Asp Thr Leu Glu Asp Asn 625 630 635 640	1920
CGT GAA TGG TAC CAG AGC ACA ATC CCT CAG AGC CCC TCT CCT GCA CCT Arg Glu Trp Tyr Gln Ser Thr Ile Pro Gln Ser Pro Ser Pro Ala Pro 645 650 655	1968
GAT GAC CCA GAG GAG GGC CGG CAG GGT CAA ACT GAG AAA TTC CAG TTT Asp Asp Pro Glu Glu Gly Arg Gln Gly Gln Thr Glu Lys Phe Gln Phe 660 665 670	2016
GAA CTA ACT TTA GAG GAA GAT GGT GAG TCA GAC ACG GAA AAG GAC AGT Glu Leu Thr Leu Glu Glu Asp Gly Glu Ser Asp Thr Glu Lys Asp Ser 675 680 685	2064
GGC AGT CAA GTG GAA GAA GAC ACT AGC TGC AGT GAC TCC AAG ACT CTT Gly Ser Gln Val Glu Glu Asp Thr Ser Cys Ser Asp Ser Lys Thr Leu 690 695 700	2112
TGT ACT CAA GAC TCA GAG TCT ACT GAA ATT CCC CTT GAT GAA CAG GTT Cys Thr Gln Asp Ser Glu Ser Thr Glu Ile Pro Leu Asp Glu Gln Val 705 710 715 720	2160
GAA GAG GAG GCA GTA GGG GAA GAA GAG GAA AGC CAG CCT GAA GCC TGT Glu Glu Glu Ala Val Gly Glu Glu Glu Glu Ser Gln Pro Glu Ala Cys 725 730 735	2208
GTC ATA GAT GAT CGT TCT CCT GAC ACG ACG GGA ATT CTG CAG TCG ACG Val Ile Asp Asp Arg Ser Pro Asp Thr Thr Gly Ile Leu Gln Ser Thr 740 745 750	2256
GTA CCG CGG GCC CGG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC Val Pro Arg Ala Arg Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly 755 760 765	2304
GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly 770 775 780	2352
GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp 785 790 795 800	2400
GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys 805 810 815	2448
CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val 820 825 830	2496
CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe	2544

835	840	845	
AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC			2592
Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe			
850	855	860	
AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC			2640
Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly			
865	870	875	880
GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG			2688
Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu			
885	890	895	
GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC			2736
Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His			
900	905	910	
AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC			2784
Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn			
915	920	925	
TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC			2832
Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp			
930	935	940	
CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC			2880
His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro			
945	950	955	960
GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC			2928
Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn			
965	970	975	
GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG			2976
Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly			
980	985	990	
ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA			3009
Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys			
995	1000		

(2) INFORMATION FOR SEQ ID NO:149:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1002 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

```

Met Ala Gln Gln Thr Ser Pro Asp Thr Leu Thr Val Pro Glu Val Asp
 1           5           10           15
Asn Pro His Cys Pro Asn Pro Trp Leu Asn Glu Asp Leu Val Lys Ser
      20           25           30
Leu Arg Glu Asn Leu Leu Gln His Glu Lys Ser Lys Thr Ala Arg Lys
      35           40           45
Ser Val Ser Pro Lys Leu Ser Pro Val Ile Ser Pro Arg Asn Ser Pro
      50           55           60
Arg Leu Leu Arg Arg Met Leu Leu Ser Ser Asn Ile Pro Lys Gln Arg
      65           70           75           80
Arg Phe Thr Val Ala His Thr Cys Phe Asp Val Asp Asn Gly Thr Ser
      85           90           95
Ala Gly Arg Ser Pro Leu Asp Pro Met Thr Ser Pro Gly Ser Gly Leu
      100           105           110
Ile Leu Gln Ala Asn Phe Val His Ser Gln Arg Arg Glu Ser Phe Leu
      115           120           125
Tyr Arg Ser Asp Ser Asp Tyr Asp Leu Ser Pro Lys Ser Met Ser Arg
      130           135           140
Asn Ser Ser Ile Ala Ser Asp Ile His Gly Asp Asp Leu Ile Val Thr
      145           150           155           160
Pro Phe Ala Gln Val Leu Ala Ser Leu Arg Thr Val Arg Asn Asn Phe
      165           170           175
Ala Ala Leu Thr Asn Leu Gln Asp Arg Ala Pro Ser Lys Arg Ser Pro
      180           185           190
Met Cys Asn Gln Pro Ser Ile Asn Lys Ala Thr Ile Thr Glu Glu Ala
      195           200           205
Tyr Gln Lys Leu Ala Ser Glu Thr Leu Glu Glu Leu Asp Trp Cys Leu
      210           215           220
Asp Gln Leu Glu Thr Leu Gln Thr Arg His Ser Val Ser Glu Met Ala
      225           230           235           240
Ser Asn Lys Phe Lys Arg Met Leu Asn Arg Glu Leu Thr His Leu Ser
      245           250           255
Glu Met Ser Arg Ser Gly Asn Gln Val Ser Glu Phe Ile Ser Asn Thr
      260           265           270
Phe Leu Asp Lys Gln His Glu Val Glu Ile Pro Ser Pro Thr Gln Lys
      275           280           285
Glu Lys Glu Lys Lys Lys Arg Pro Met Ser Gln Ile Ser Gly Val Lys
      290           295           300
Lys Leu Met His Ser Ser Ser Leu Thr Asn Ser Ser Ile Pro Arg Phe
      305           310           315           320
Gly Val Lys Thr Glu Gln Glu Asp Val Leu Ala Lys Glu Leu Glu Asp
      325           330           335
Val Asn Lys Trp Gly Leu His Val Phe Arg Ile Ala Glu Leu Ser Gly
      340           345           350
Asn Arg Pro Leu Thr Val Ile Met His Thr Ile Phe Gln Glu Arg Asp
      355           360           365
Leu Leu Lys Thr Phe Lys Ile Pro Val Asp Thr Leu Ile Thr Tyr Leu
      370           375           380
Met Thr Leu Glu Asp His Tyr His Ala Asp Val Ala Tyr His Asn Asn
      385           390           395           400
Ile His Ala Ala Asp Val Val Gln Ser Thr His Val Leu Leu Ser Thr
      405           410           415

```


Pro Ala Leu Glu Ala Val Phe Thr Asp Leu Glu Ile Leu Ala Ala Ile
 420 425 430
 Phe Ala Ser Ala Ile His Asp Val Asp His Pro Gly Val Ser Asn Gln
 435 440 445
 Phe Leu Ile Asn Thr Asn Ser Glu Leu Ala Leu Met Tyr Asn Asp Ser
 450 455 460
 Ser Val Leu Glu Asn His His Leu Ala Val Gly Phe Lys Leu Leu Gln
 465 470 475 480
 Glu Glu Asn Cys Asp Ile Phe Gln Asn Leu Thr Lys Lys Gln Arg Gln
 485 490 495
 Ser Leu Arg Lys Met Val Ile Asp Ile Val Leu Ala Thr Asp Met Ser
 500 505 510
 Lys His Met Asn Leu Leu Ala Asp Leu Lys Thr Met Val Glu Thr Lys
 515 520 525
 Lys Val Thr Ser Ser Gly Val Leu Leu Leu Asp Asn Tyr Ser Asp Arg
 530 535 540
 Ile Gln Val Leu Gln Asn Met Val His Cys Ala Asp Leu Ser Asn Pro
 545 550 555 560
 Thr Lys Pro Leu Gln Leu Tyr Arg Gln Trp Thr Asp Arg Ile Met Glu
 565 570 575
 Glu Phe Phe Arg Gln Gly Asp Arg Glu Arg Glu Arg Gly Met Glu Ile
 580 585 590
 Ser Pro Met Cys Asp Lys His Asn Ala Ser Val Glu Lys Ser Gln Val
 595 600 605
 Gly Phe Ile Asp Tyr Ile Val His Pro Leu Trp Glu Thr Trp Ala Asp
 610 615 620
 Leu Val His Pro Asp Ala Gln Asp Ile Leu Asp Thr Leu Glu Asp Asn
 625 630 635 640
 Arg Glu Trp Tyr Gln Ser Thr Ile Pro Gln Ser Pro Ser Pro Ala Pro
 645 650 655
 Asp Asp Pro Glu Glu Gly Arg Gln Gly Gln Thr Glu Lys Phe Gln Phe
 660 665 670
 Glu Leu Thr Leu Glu Glu Asp Gly Glu Ser Asp Thr Glu Lys Asp Ser
 675 680 685
 Gly Ser Gln Val Glu Glu Asp Thr Ser Cys Ser Asp Ser Lys Thr Leu
 690 695 700
 Cys Thr Gln Asp Ser Glu Ser Thr Glu Ile Pro Leu Asp Glu Gln Val
 705 710 715 720
 Glu Glu Glu Ala Val Gly Glu Glu Glu Glu Ser Gln Pro Glu Ala Cys
 725 730 735
 Val Ile Asp Asp Arg Ser Pro Asp Thr Thr Gly Ile Leu Gln Ser Thr
 740 745 750
 Val Pro Arg Ala Arg Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly
 755 760 765
 Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly
 770 775 780
 Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp
 785 790 795 800
 Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys
 805 810 815
 Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val
 820 825 830
 Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe
 835 840 845

Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe
 850 855 860
 Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly
 865 870 875 880
 Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu
 885 890 895
 Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His
 900 905 910
 Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn
 915 920 925
 Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp
 930 935 940
 His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro
 945 950 955 960
 Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn
 965 970 975
 Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly
 980 985 990
 Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
 995 1000

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3201 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...3198
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

ATG GAG GCA GAG GGC AGC AGC GCG CCG GCC CGG GCG GGC AGC GGA GAG	48
Met Glu Ala Glu Gly Ser Ser Ala Pro Ala Arg Ala Gly Ser Gly Glu	
1 5 10 15	
GGC AGC GAC AGC GCC GGC GGG GCC ACG CTC AAA GCC CCC AAG CAT CTC	96
Gly Ser Asp Ser Ala Gly Gly Ala Thr Leu Lys Ala Pro Lys His Leu	
20 25 30	
TGG AGG CAC GAG CAG CAC CAC CAG TAC CCG CTC CGG CAG CCC CAG TTC	144
Trp Arg His Glu Gln His His Gln Tyr Pro Leu Arg Gln Pro Gln Phe	
35 40 45	
CGC CTC CTG CAT CCC CAT CAC CAC CTG CCC CCG CCG CCG CCA CCC TCG	192
Arg Leu Leu His Pro His His His Leu Pro Pro Pro Pro Pro Ser	
50 55 60	

CCC CAG CCC CAG CCC CAG TGT CCG CTA CAG CCG CCG CCG CCG CCC CCC	240
Pro Gln Pro Gln Pro Gln Cys Pro Leu Gln Pro Pro Pro Pro Pro Pro	
65 70 75 80	
CTG CCG CCG CCC CCG CCG CCG CCC GGG GCT GCC CGC GGC CGC TAC GCC	288
Leu Pro Pro Pro Pro Pro Pro Pro Gly Ala Ala Arg Gly Arg Tyr Ala	
85 90 95	
TCG AGC GGG GCC ACC GGC CGC GTC CCG CAT CGC GGC TAC TCG GAC ACC	336
Ser Ser Gly Ala Thr Gly Arg Val Arg His Arg Gly Tyr Ser Asp Thr	
100 105 110	
GAG CGC TAC CTG TAC TGT CGC GCC ATG GAC CGC ACC TCC TAC GCG GTG	384
Glu Arg Tyr Leu Tyr Cys Arg Ala Met Asp Arg Thr Ser Tyr Ala Val	
115 120 125	
GAG ACC GGC CAC CGG CCC GGC CTG AAG AAA TCC AGG ATG TCC TGG CCC	432
Glu Thr Gly His Arg Pro Gly Leu Lys Lys Ser Arg Met Ser Trp Pro	
130 135 140	
TCC TCG TTC CAG GGA CTC AGG CGT TTT GAT GTG GAC AAT GGC ACA TCT	480
Ser Ser Phe Gln Gly Leu Arg Arg Phe Asp Val Asp Asn Gly Thr Ser	
145 150 155 160	
GCG GGA CGG AGT CCC TTG GAT CCC ATG ACC AGC CCA GGA TCC GGG CTA	528
Ala Gly Arg Ser Pro Leu Asp Pro Met Thr Ser Pro Gly Ser Gly Leu	
165 170 175	
ATT CTC CAA GCA AAT TTT GTC CAC AGT CAA CGA CGG GAG TCC TTC CTG	576
Ile Leu Gln Ala Asn Phe Val His Ser Gln Arg Arg Glu Ser Phe Leu	
180 185 190	
TAT CGA TCC GAC AGC GAT TAT GAC CTC TCT CCA AAG TCT ATG TCC CGG	624
Tyr Arg Ser Asp Ser Asp Tyr Asp Leu Ser Pro Lys Ser Met Ser Arg	
195 200 205	
AAC TCC TCC ATT GCC AGT GAT ATA CAC GGA GAT GAC TTG ATT GTG ACT	672
Asn Ser Ser Ile Ala Ser Asp Ile His Gly Asp Asp Leu Ile Val Thr	
210 215 220	
CCA TTT GCT CAG GTC TTG GCC AGT CTG CGA ACT GTA CGA AAC AAC TTT	720
Pro Phe Ala Gln Val Leu Ala Ser Leu Arg Thr Val Arg Asn Asn Phe	
225 230 235 240	
GCT GCA TTA ACT AAT TTG CAA GAT CGA GCA CCT AGC AAA AGA TCA CCC	768
Ala Ala Leu Thr Asn Leu Gln Asp Arg Ala Pro Ser Lys Arg Ser Pro	
245 250 255	
ATG TGC AAC CAA CCA TCC ATC AAC AAA GCC ACC ATA ACA GAG GAG GCC	816
Met Cys Asn Gln Pro Ser Ile Asn Lys Ala Thr Ile Thr Glu Glu Ala	
260 265 270	
TAC CAG AAA CTG GCC AGC GAG ACC CTG GAG GAG CTG GAC TGG TGT CTG	864
Tyr Gln Lys Leu Ala Ser Glu Thr Leu Glu Glu Leu Asp Trp Cys Leu	

275	280	285	
GAC CAG CTA GAG ACC CTA CAG ACC AGG CAC TCC GTC AGT GAG ATG GCC Asp Gln Leu Glu Thr Leu Gln Thr Arg His Ser Val Ser Glu Met Ala 290 295 300			912
TCC AAC AAG TTT AAA AGG ATG CTT AAT CGG GAG CTC ACC CAT CTC TCT Ser Asn Lys Phe Lys Arg Met Leu Asn Arg Glu Leu Thr His Leu Ser 305 310 315 320			960
GAA ATG AGT CGG TCT GGA AAT CAA GTG TCA GAG TTT ATA TCA AAC ACA Glu Met Ser Arg Ser Gly Asn Gln Val Ser Glu Phe Ile Ser Asn Thr 325 330 335			1008
TTC TTA GAT AAG CAA CAT GAA GTG GAA ATT CCT TCT CCA ACT CAG AAG Phe Leu Asp Lys Gln His Glu Val Glu Ile Pro Ser Pro Thr Gln Lys 340 345 350			1056
GAA AAG GAG AAA AAG AAA AGA CCA ATG TCT CAG ATC AGT GGA GTC AAG Glu Lys Glu Lys Lys Lys Arg Pro Met Ser Gln Ile Ser Gly Val Lys 355 360 365			1104
AAA TTG ATG CAC AGC TCT AGT CTG ACT AAT TCA AGT ATC CCA AGG TTT Lys Leu Met His Ser Ser Ser Leu Thr Asn Ser Ser Ile Pro Arg Phe 370 375 380			1152
GGA GTT AAA ACT GAA CAA GAA GAT GTC CTT GCC AAG GAA CTA GAA GAT Gly Val Lys Thr Glu Gln Glu Asp Val Leu Ala Lys Glu Leu Glu Asp 385 390 395 400			1200
GTG AAC AAA TGG GGT CTT CAT GTT TTC AGA ATA GCA GAG TTG TCT GGT Val Asn Lys Trp Gly Leu His Val Phe Arg Ile Ala Glu Leu Ser Gly 405 410 415			1248
AAC CGG CCC TTG ACT GTT ATC ATG CAC ACC ATT TTT CAG GAA CGG GAT Asn Arg Pro Leu Thr Val Ile Met His Thr Ile Phe Gln Glu Arg Asp 420 425 430			1296
TTA TTA AAA ACA TTT AAA ATT CCA GTA GAT ACT TTA ATT ACA TAT CTT Leu Leu Lys Thr Phe Lys Ile Pro Val Asp Thr Leu Ile Thr Tyr Leu 435 440 445			1344
ATG ACT CTC GAA GAC CAT TAC CAT GCT GAT GTG GCC TAT CAC AAC AAT Met Thr Leu Glu Asp His Tyr His Ala Asp Val Ala Tyr His Asn Asn 450 455 460			1392
ATC CAT GCT GCA GAT GTT GTC CAG TCT ACT CAT GTG CTA TTA TCT ACA Ile His Ala Ala Asp Val Val Gln Ser Thr His Val Leu Leu Ser Thr 465 470 475 480			1440
CCT GCT TTG GAG GCT GTG TTT ACA GAT TTG GAG ATT CTT GCA GCA ATT Pro Ala Leu Glu Ala Val Phe Thr Asp Leu Glu Ile Leu Ala Ala Ile 485 490 495			1488

TTT GCC AGT GCA ATA CAT GAT GTA GAT CAT CCT GGT GTG TCC AAT CAA	1536
Phe Ala Ser Ala Ile His Asp Val Asp His Pro Gly Val Ser Asn Gln	
500 505 510	
TTT CTG ATC AAT ACA AAC TCT GAA CTT GCC TTG ATG TAC AAT GAT TCC	1584
Phe Leu Ile Asn Thr Asn Ser Glu Leu Ala Leu Met Tyr Asn Asp Ser	
515 520 525	
TCA GTC TTA GAG AAC CAT CAT TTG GCT GTG GGC TTT AAA TTG CTT CAG	1632
Ser Val Leu Glu Asn His His Leu Ala Val Gly Phe Lys Leu Leu Gln	
530 535 540	
GAA GAA AAC TGT GAC ATT TTC CAG AAT TTG ACC AAA AAA CAA AGA CAA	1680
Glu Glu Asn Cys Asp Ile Phe Gln Asn Leu Thr Lys Lys Gln Arg Gln	
545 550 555 560	
TCT TTA AGG AAA ATG GTC ATT GAC ATC GTA CTT GCA ACA GAT ATG TCA	1728
Ser Leu Arg Lys Met Val Ile Asp Ile Val Leu Ala Thr Asp Met Ser	
565 570 575	
AAA CAC ATG AAT CTA CTG GCT GAT TTG AAG ACT ATG GTT GAA ACT AAG	1776
Lys His Met Asn Leu Leu Ala Asp Leu Lys Thr Met Val Glu Thr Lys	
580 585 590	
AAA GTG ACA AGC TCT GGA GTT CTT CTT CTT GAT AAT TAT TCC GAT AGG	1824
Lys Val Thr Ser Ser Gly Val Leu Leu Leu Asp Asn Tyr Ser Asp Arg	
595 600 605	
ATT CAG GTT CTT CAG AAT ATG GTG CAC TGT GCA GAT CTG AGC AAC CCA	1872
Ile Gln Val Leu Gln Asn Met Val His Cys Ala Asp Leu Ser Asn Pro	
610 615 620	
ACA AAG CCT CTC CAG CTG TAC CGC CAG TGG ACG GAC CGG ATA ATG GAG	1920
Thr Lys Pro Leu Gln Leu Tyr Arg Gln Trp Thr Asp Arg Ile Met Glu	
625 630 635 640	
GAG TTC TTC CGC CAA GGA GAC CGA GAG AGG GAA CGT GGC ATG GAG ATA	1968
Glu Phe Phe Arg Gln Gly Asp Arg Glu Arg Glu Arg Gly Met Glu Ile	
645 650 655	
AGC CCC ATG TGT GAC AAG CAC AAT GCT TCC GTG GAA AAA TCA CAG GTG	2016
Ser Pro Met Cys Asp Lys His Asn Ala Ser Val Glu Lys Ser Gln Val	
660 665 670	
GGC TTC ATA GAC TAT ATT GTT CAT CCC CTC TGG GAG ACA TGG GCA GAC	2064
Gly Phe Ile Asp Tyr Ile Val His Pro Leu Trp Glu Thr Trp Ala Asp	
675 680 685	
CTC GTC CAC CCT GAC GCC CAG GAT ATT TTG GAC ACT TTG GAG GAC AAT	2112
Leu Val His Pro Asp Ala Gln Asp Ile Leu Asp Thr Leu Glu Asp Asn	
690 695 700	
CGT GAA TGG TAC CAG AGC ACA ATC CCT CAG AGC CCC TCT CCT GCA CCT	2160
Arg Glu Trp Tyr Gln Ser Thr Ile Pro Gln Ser Pro Ser Pro Ala Pro	

705	710	715	720	
GAT GAC CCA GAG GAG GGC CGG CAG GGT CAA ACT GAG AAA TTC CAG TTT				2208
Asp Asp Pro Glu Glu Gly Arg Gln Gly Gln Thr Glu Lys Phe Gln Phe				
725	730	735		
GAA CTA ACT TTA GAG GAA GAT GGT GAG TCA GAC ACG GAA AAG GAC AGT				2256
Glu Leu Thr Leu Glu Glu Asp Gly Glu Ser Asp Thr Glu Lys Asp Ser				
740	745	750		
GGC AGT CAA GTG GAA GAA GAC ACT AGC TGC AGT GAC TCC AAG ACT CTT				2304
Gly Ser Gln Val Glu Glu Asp Thr Ser Cys Ser Asp Ser Lys Thr Leu				
755	760	765		
TGT ACT CAA GAC TCA GAG TCT ACT GAA ATT CCC CTT GAT GAA CAG GTT				2352
Cys Thr Gln Asp Ser Glu Ser Thr Glu Ile Pro Leu Asp Glu Gln Val				
770	775	780		
GAA GAG GAG GCA GTA GGG GAA GAA GAG GAA AGC CAG CCT GAA GCC TGT				2400
Glu Glu Glu Ala Val Gly Glu Glu Glu Glu Ser Gln Pro Glu Ala Cys				
785	790	795	800	
GTC ATA GAT GAT CGT TCT CCT GAC ACG ACG GGA ATT CTG CAG TCG ACG				2448
Val Ile Asp Asp Arg Ser Pro Asp Thr Thr Gly Ile Leu Gln Ser Thr				
805	810	815		
GTA CCG CGG GCC CGG GAT CCA CCG GTC GCC ACC ATG GTG AGC AAG GGC				2496
Val Pro Arg Ala Arg Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly				
820	825	830		
GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC				2544
Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly				
835	840	845		
GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT				2592
Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp				
850	855	860		
GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG				2640
Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys				
865	870	875	880	
CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG				2688
Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val				
885	890	895		
CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC				2736
Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe				
900	905	910		
AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC				2784
Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe				
915	920	925		

AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC	2832
Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly	
930 935 940	
GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG	2880
Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu	
945 950 955 960	
GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC	2928
Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His	
965 970 975	
AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC	2976
Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn	
980 985 990	
TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC	3024
Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp	
995 1000 1005	
CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC	3072
His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro	
1010 1015 1020	
GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC	3120
Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn	
1025 1030 1035 1040	
GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG	3168
Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly	
1045 1050 1055	
ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA	3201
Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys	
1060 1065	

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1066 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

Met Glu Ala Glu Gly Ser Ser Ala Pro Ala Arg Ala Gly Ser Gly Glu	
1 5 10 15	
Gly Ser Asp Ser Ala Gly Gly Ala Thr Leu Lys Ala Pro Lys His Leu	
20 25 30	

Trp Arg His Glu Gln His His Gln Tyr Pro Leu Arg Gln Pro Gln Phe
 35 40 45
 Arg Leu Leu His Pro His His His Leu Pro Pro Pro Pro Pro Pro Ser
 50 55 60
 Pro Gln Pro Gln Pro Gln Cys Pro Leu Gln Pro Pro Pro Pro Pro Pro
 65 70 75 80
 Leu Pro Pro Pro Pro Pro Pro Pro Gly Ala Ala Arg Gly Arg Tyr Ala
 85 90 95
 Ser Ser Gly Ala Thr Gly Arg Val Arg His Arg Gly Tyr Ser Asp Thr
 100 105 110
 Glu Arg Tyr Leu Tyr Cys Arg Ala Met Asp Arg Thr Ser Tyr Ala Val
 115 120 125
 Glu Thr Gly His Arg Pro Gly Leu Lys Lys Ser Arg Met Ser Trp Pro
 130 135 140
 Ser Ser Phe Gln Gly Leu Arg Arg Phe Asp Val Asp Asn Gly Thr Ser
 145 150 155 160
 Ala Gly Arg Ser Pro Leu Asp Pro Met Thr Ser Pro Gly Ser Gly Leu
 165 170 175
 Ile Leu Gln Ala Asn Phe Val His Ser Gln Arg Arg Glu Ser Phe Leu
 180 185 190
 Tyr Arg Ser Asp Ser Asp Tyr Asp Leu Ser Pro Lys Ser Met Ser Arg
 195 200 205
 Asn Ser Ser Ile Ala Ser Asp Ile His Gly Asp Asp Leu Ile Val Thr
 210 215 220
 Pro Phe Ala Gln Val Leu Ala Ser Leu Arg Thr Val Arg Asn Asn Phe
 225 230 235 240
 Ala Ala Leu Thr Asn Leu Gln Asp Arg Ala Pro Ser Lys Arg Ser Pro
 245 250 255
 Met Cys Asn Gln Pro Ser Ile Asn Lys Ala Thr Ile Thr Glu Glu Ala
 260 265 270
 Tyr Gln Lys Leu Ala Ser Glu Thr Leu Glu Glu Leu Asp Trp Cys Leu
 275 280 285
 Asp Gln Leu Glu Thr Leu Gln Thr Arg His Ser Val Ser Glu Met Ala
 290 295 300
 Ser Asn Lys Phe Lys Arg Met Leu Asn Arg Glu Leu Thr His Leu Ser
 305 310 315 320
 Glu Met Ser Arg Ser Gly Asn Gln Val Ser Glu Phe Ile Ser Asn Thr
 325 330 335
 Phe Leu Asp Lys Gln His Glu Val Glu Ile Pro Ser Pro Thr Gln Lys
 340 345 350
 Glu Lys Glu Lys Lys Lys Arg Pro Met Ser Gln Ile Ser Gly Val Lys
 355 360 365
 Lys Leu Met His Ser Ser Ser Leu Thr Asn Ser Ser Ile Pro Arg Phe
 370 375 380
 Gly Val Lys Thr Glu Gln Glu Asp Val Leu Ala Lys Glu Leu Glu Asp
 385 390 395 400
 Val Asn Lys Trp Gly Leu His Val Phe Arg Ile Ala Glu Leu Ser Gly
 405 410 415
 Asn Arg Pro Leu Thr Val Ile Met His Thr Ile Phe Gln Glu Arg Asp
 420 425 430
 Leu Leu Lys Thr Phe Lys Ile Pro Val Asp Thr Leu Ile Thr Tyr Leu
 435 440 445
 Met Thr Leu Glu Asp His Tyr His Ala Asp Val Ala Tyr His Asn Asn
 450 455 460

Ile	His	Ala	Ala	Asp	Val	Val	Gln	Ser	Thr	His	Val	Leu	Leu	Ser	Thr	465	470	475	480
Pro	Ala	Leu	Glu	Ala	Val	Phe	Thr	Asp	Leu	Glu	Ile	Leu	Ala	Ala	Ile	485	490	495	
Phe	Ala	Ser	Ala	Ile	His	Asp	Val	Asp	His	Pro	Gly	Val	Ser	Asn	Gln	500	505	510	
Phe	Leu	Ile	Asn	Thr	Asn	Ser	Glu	Leu	Ala	Leu	Met	Tyr	Asn	Asp	Ser	515	520	525	
Ser	Val	Leu	Glu	Asn	His	His	Leu	Ala	Val	Gly	Phe	Lys	Leu	Leu	Gln	530	535	540	
Glu	Glu	Asn	Cys	Asp	Ile	Phe	Gln	Asn	Leu	Thr	Lys	Lys	Gln	Arg	Gln	545	550	555	560
Ser	Leu	Arg	Lys	Met	Val	Ile	Asp	Ile	Val	Leu	Ala	Thr	Asp	Met	Ser	565	570	575	
Lys	His	Met	Asn	Leu	Leu	Ala	Asp	Leu	Lys	Thr	Met	Val	Glu	Thr	Lys	580	585	590	
Lys	Val	Thr	Ser	Ser	Gly	Val	Leu	Leu	Leu	Asp	Asn	Tyr	Ser	Asp	Arg	595	600	605	
Ile	Gln	Val	Leu	Gln	Asn	Met	Val	His	Cys	Ala	Asp	Leu	Ser	Asn	Pro	610	615	620	
Thr	Lys	Pro	Leu	Gln	Leu	Tyr	Arg	Gln	Trp	Thr	Asp	Arg	Ile	Met	Glu	625	630	635	640
Glu	Phe	Phe	Arg	Gln	Gly	Asp	Arg	Glu	Arg	Glu	Arg	Gly	Met	Glu	Ile	645	650	655	
Ser	Pro	Met	Cys	Asp	Lys	His	Asn	Ala	Ser	Val	Glu	Lys	Ser	Gln	Val	660	665	670	
Gly	Phe	Ile	Asp	Tyr	Ile	Val	His	Pro	Leu	Trp	Glu	Thr	Trp	Ala	Asp	675	680	685	
Leu	Val	His	Pro	Asp	Ala	Gln	Asp	Ile	Leu	Asp	Thr	Leu	Glu	Asp	Asn	690	695	700	
Arg	Glu	Trp	Tyr	Gln	Ser	Thr	Ile	Pro	Gln	Ser	Pro	Ser	Pro	Ala	Pro	705	710	715	720
Asp	Asp	Pro	Glu	Glu	Gly	Arg	Gln	Gly	Gln	Thr	Glu	Lys	Phe	Gln	Phe	725	730	735	
Glu	Leu	Thr	Leu	Glu	Glu	Asp	Gly	Glu	Ser	Asp	Thr	Glu	Lys	Asp	Ser	740	745	750	
Gly	Ser	Gln	Val	Glu	Glu	Asp	Thr	Ser	Cys	Ser	Asp	Ser	Lys	Thr	Leu	755	760	765	
Cys	Thr	Gln	Asp	Ser	Glu	Ser	Thr	Glu	Ile	Pro	Leu	Asp	Glu	Gln	Val	770	775	780	
Glu	Glu	Glu	Ala	Val	Gly	Glu	Glu	Glu	Glu	Ser	Gln	Pro	Glu	Ala	Cys	785	790	795	800
Val	Ile	Asp	Asp	Arg	Ser	Pro	Asp	Thr	Thr	Gly	Ile	Leu	Gln	Ser	Thr	805	810	815	
Val	Pro	Arg	Ala	Arg	Asp	Pro	Pro	Val	Ala	Thr	Met	Val	Ser	Lys	Gly	820	825	830	
Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly	835	840	845	
Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp	850	855	860	
Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys	865	870	875	880
Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val	885	890	895	

Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe
 900 905 910
 Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe
 915 920 925
 Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly
 930 935 940
 Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu
 945 950 955 960
 Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His
 965 970 975
 Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn
 980 985 990
 Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp
 995 1000 1005
 His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro
 1010 1015 1020
 Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn
 1025 1030 1035 1040
 Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly
 1045 1050 1055
 Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
 1060 1065

(2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3024 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...3021
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

ATG AGC TGG TCA CCT TCC CTG ACA ACG CAG ACA TGT GGG GCC TGG GAA	48
Met Ser Trp Ser Pro Ser Leu Thr Thr Gln Thr Cys Gly Ala Trp Glu	
1 5 10 15	
ATG AAA GAG CGC CTT GGG ACA GGG GGA TTT GGA AAT GTC ATC CGA TGG	96
Met Lys Glu Arg Leu Gly Thr Gly Gly Phe Gly Asn Val Ile Arg Trp	
20 25 30	
CAC AAT CAG GAA ACA GGT GAG CAG ATT GCC ATC AAG CAG TGC CGG CAG	144
His Asn Gln Glu Thr Gly Glu Gln Ile Ala Ile Lys Gln Cys Arg Gln	
35 40 45	
GAG CTC AGC CCC CGG AAC CGA GAG CGG TGG TGC CTG GAG ATC CAG ATC	192
Glu Leu Ser Pro Arg Asn Arg Glu Arg Trp Cys Leu Glu Ile Gln Ile	

50	55	60	
ATG AGA AGG CTG ACC CAC CCC AAT GTG GTG GCT GCC CGA GAT GTC CCT			240
Met Arg Arg Leu Thr His Pro Asn Val Val Ala Ala Arg Asp Val Pro			
65	70	75	80
GAG GGG ATG CAG AAC TTG GCG CCC AAT GAC CTG CCC CTG CTG GCC ATG			288
Glu Gly Met Gln Asn Leu Ala Pro Asn Asp Leu Pro Leu Leu Ala Met			
	85	90	95
GAG TAC TGC CAA GGA GGA GAT CTC CGG AAG TAC CTG AAC CAG TTT GAG			336
Glu Tyr Cys Gln Gly Gly Asp Leu Arg Lys Tyr Leu Asn Gln Phe Glu			
	100	105	110
AAC TGC TGT GGT CTG CGG GAA GGT GCC ATC CTC ACC TTG CTG AGT GAC			384
Asn Cys Cys Gly Leu Arg Glu Gly Ala Ile Leu Thr Leu Leu Ser Asp			
	115	120	125
ATT GCC TCT GCG CTT AGA TAC CTT CAT GAA AAC AGA ATC ATC CAT CGG			432
Ile Ala Ser Ala Leu Arg Tyr Leu His Glu Asn Arg Ile Ile His Arg			
	130	135	140
GAT CTA AAG CCA GAA AAC ATC GTC CTG CAG CAA GGA GAA CAG AGG TTA			480
Asp Leu Lys Pro Glu Asn Ile Val Leu Gln Gln Gly Glu Gln Arg Leu			
145	150	155	160
ATA CAC AAA ATT ATT GAC CTA GGA TAT GCC AAG GAG CTG GAT CAG GGC			528
Ile His Lys Ile Ile Asp Leu Gly Tyr Ala Lys Glu Leu Asp Gln Gly			
	165	170	175
AGT CTT TGC ACA TCA TTC GTG GGG ACC CTG CAG TAC CTG GCC CCA GAG			576
Ser Leu Cys Thr Ser Phe Val Gly Thr Leu Gln Tyr Leu Ala Pro Glu			
	180	185	190
CTA CTG GAG CAG CAG AAG TAC ACA GTG ACC GTC GAC TAC TGG AGC TTC			624
Leu Leu Glu Gln Gln Lys Tyr Thr Val Thr Val Asp Tyr Trp Ser Phe			
	195	200	205
GGC ACC CTG GCC TTT GAG TGC ATC ACG GGC TTC CGG CCC TTC CTC CCC			672
Gly Thr Leu Ala Phe Glu Cys Ile Thr Gly Phe Arg Pro Phe Leu Pro			
210	215	220	
AAC TGG CAG CCC GTG CAG TGG CAT TCA AAA GTG CGG CAG AAG AGT GAG			720
Asn Trp Gln Pro Val Gln Trp His Ser Lys Val Arg Gln Lys Ser Glu			
225	230	235	240
GTG GAC ATT GTT GTT AGC GAA GAC TTG AAT GGA ACG GTG AAG TTT TCA			768
Val Asp Ile Val Val Ser Glu Asp Leu Asn Gly Thr Val Lys Phe Ser			
	245	250	255
AGC TCT TTA CCC TAC CCC AAT AAT CTT AAC AGT GTC CTG GCT GAG CGA			816
Ser Ser Leu Pro Tyr Pro Asn Asn Leu Asn Ser Val Leu Ala Glu Arg			
	260	265	270

CTG GAG AAG TGG CTG CAA CTG ATG CTG ATG TGG CAC CCC CGA CAG AGG Leu Glu Lys Trp Leu Gln Leu Met Leu Met Trp His Pro Arg Gln Arg 275 280 285	864
GGC ACG GAT CCC ACG TAT GGG CCC AAT GGC TGC TTC AAG GCC CTG GAT Gly Thr Asp Pro Thr Tyr Gly Pro Asn Gly Cys Phe Lys Ala Leu Asp 290 295 300	912
GAC ATC TTA AAC TTA AAG CTG GTT CAT ATC TTG AAC ATG GTC ACG GGC Asp Ile Leu Asn Leu Lys Leu Val His Ile Leu Asn Met Val Thr Gly 305 310 315 320	960
ACC ATC CAC ACC TAC CCT GTG ACA GAG GAT GAG AGT CTG CAG AGC TTG Thr Ile His Thr Tyr Pro Val Thr Glu Asp Glu Ser Leu Gln Ser Leu 325 330 335	1008
AAG GCC AGA ATC CAA CAG GAC ACG GGC ATC CCA GAG GAG GAC CAG GAG Lys Ala Arg Ile Gln Gln Asp Thr Gly Ile Pro Glu Glu Asp Gln Glu 340 345 350	1056
CTG CTG CAG GAA GCG GGC CTG GCG TTG ATC CCC GAT AAG CCT GCC ACT Leu Leu Gln Glu Ala Gly Leu Ala Leu Ile Pro Asp Lys Pro Ala Thr 355 360 365	1104
CAG TGT ATT TCA GAC GGC AAG TTA AAT GAG GGC CAC ACA TTG GAC ATG Gln Cys Ile Ser Asp Gly Lys Leu Asn Glu Gly His Thr Leu Asp Met 370 375 380	1152
GAT CTT GTT TTT CTC TTT GAC AAC AGT AAA ATC ACC TAT GAG ACT CAG Asp Leu Val Phe Leu Phe Asp Asn Ser Lys Ile Thr Tyr Glu Thr Gln 385 390 395 400	1200
ATC TCC CCA CGG CCC CAA CCT GAA AGT GTC AGC TGT ATC CTT CAA GAG Ile Ser Pro Arg Pro Gln Pro Glu Ser Val Ser Cys Ile Leu Gln Glu 405 410 415	1248
CCC AAG AGG AAT CTC GCC TTC TTC CAG CTG AGG AAG GTG TGG GGC CAG Pro Lys Arg Asn Leu Ala Phe Phe Gln Leu Arg Lys Val Trp Gly Gln 420 425 430	1296
GTC TGG CAC AGC ATC CAG ACC CTG AAG GAA GAT TGC AAC CGG CTG CAG Val Trp His Ser Ile Gln Thr Leu Lys Glu Asp Cys Asn Arg Leu Gln 435 440 445	1344
CAG GGA CAG CGA GCC GCC ATG ATG AAT CTC CTC CGA AAC AAC AGC TGC Gln Gly Gln Arg Ala Ala Met Met Asn Leu Leu Arg Asn Asn Ser Cys 450 455 460	1392
CTC TCC AAA ATG AAG AAT TCC ATG GCT TCC ATG TCT CAG CAG CTC AAG Leu Ser Lys Met Lys Asn Ser Met Ala Ser Met Ser Gln Gln Leu Lys 465 470 475 480	1440
GCC AAG TTG GAT TTC TTC AAA ACC AGC ATC CAG ATT GAC CTG GAG AAG Ala Lys Leu Asp Phe Phe Lys Thr Ser Ile Gln Ile Asp Leu Glu Lys	1488

485	490	495	
TAC AGC GAG CAA ACC GAG TTT GGG ATC ACA TCA GAT AAA CTG CTG CTG			1536
Tyr Ser Glu Gln Thr Glu Phe Gly Ile Thr Ser Asp Lys Leu Leu Leu			
500	505	510	
GCC TGG AGG GAA ATG GAG CAG GCT GTG GAG CTC TGT GGG CGG GAG AAC			1584
Ala Trp Arg Glu Met Glu Gln Ala Val Glu Leu Cys Gly Arg Glu Asn			
515	520	525	
GAA GTG AAA CTC CTG GTA GAA CGG ATG ATG GCT CTG CAG ACC GAC ATT			1632
Glu Val Lys Leu Leu Val Glu Arg Met Met Ala Leu Gln Thr Asp Ile			
530	535	540	
GTG GAC TTA CAG AGG AGC CCC ATG GGC CGG AAG CAG GGG GGA ACG CTG			1680
Val Asp Leu Gln Arg Ser Pro Met Gly Arg Lys Gln Gly Gly Thr Leu			
545	550	555	560
GAC GAC CTA GAG GAG CAA GCA AGG GAG CTG TAC AGG AGA CTA AGG GAA			1728
Asp Asp Leu Glu Gln Ala Arg Glu Leu Tyr Arg Arg Leu Arg Glu			
565	570	575	
AAA CCT CGA GAC CAG CGA ACT GAG GGT GAC AGT CAG GAA ATG GTA CGG			1776
Lys Pro Arg Asp Gln Arg Thr Glu Gly Asp Ser Gln Glu Met Val Arg			
580	585	590	
CTG CTG CTT CAG GCA ATT CAG AGC TTC GAG AAG AAA GTG CGA GTG ATC			1824
Leu Leu Leu Gln Ala Ile Gln Ser Phe Glu Lys Lys Val Arg Val Ile			
595	600	605	
TAT ACG CAG CTC AGT AAA ACT GTG GTT TGC AAG CAG AAG GCG CTG GAA			1872
Tyr Thr Gln Leu Ser Lys Thr Val Val Cys Lys Gln Lys Ala Leu Glu			
610	615	620	
CTG TTG CCC AAG GTG GAA GAG GTG GTG AGC TTA ATG AAT GAG GAT GAG			1920
Leu Leu Pro Lys Val Glu Glu Val Val Ser Leu Met Asn Glu Asp Glu			
625	630	635	640
AAG ACT GTT GTC CGG CTG CAG GAG AAG CGG CAG AAG GAG CTC TGG AAT			1968
Lys Thr Val Val Arg Leu Gln Glu Lys Arg Gln Lys Glu Leu Trp Asn			
645	650	655	
CTC CTG AAG ATT GCT TGT AGC AAG GTC CGT GGT CCT GTC AGT GGA AGC			2016
Leu Leu Lys Ile Ala Cys Ser Lys Val Arg Gly Pro Val Ser Gly Ser			
660	665	670	
CCG GAT AGC ATG AAT GCC TCT CGA CTT AGC CAG CCT GGG CAG CTG ATG			2064
Pro Asp Ser Met Asn Ala Ser Arg Leu Ser Gln Pro Gly Gln Leu Met			
675	680	685	
TCT CAG CCC TCC ACG GCC TCC AAC AGC TTA CCT GAG CCA GCC AAG AAG			2112
Ser Gln Pro Ser Thr Ala Ser Asn Ser Leu Pro Glu Pro Ala Lys Lys			
690	695	700	

AGT GAA GAA CTG GTG GCT GAA GCA CAT AAC CTC TGC ACC CTG CTA GAA Ser Glu Glu Leu Val Ala Glu Ala His Asn Leu Cys Thr Leu Leu Glu 705 710 715 720	2160
AAT GCC ATA CAG GAC ACT GTG AGG GAA CAA GAC CAG AGT TTC ACG GCC Asn Ala Ile Gln Asp Thr Val Arg Glu Gln Asp Gln Ser Phe Thr Ala 725 730 735	2208
CTA GAC TGG AGC TGG TTA CAG ACG GAA GAA GAA GAG CAC AGC TGC CTG Leu Asp Trp Ser Trp Leu Gln Thr Glu Glu Glu Glu His Ser Cys Leu 740 745 750	2256
GAG CAG GCC TCA TGG GTA CCG CGG GCC CGG GAT CCA CCG GTC GCC ACC Glu Gln Ala Ser Trp Val Pro Arg Ala Arg Asp Pro Pro Val Ala Thr 755 760 765	2304
ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG CCC ATC CTG Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu 770 775 780	2352
GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC GTG TCC GGC Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly 785 790 795 800	2400
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile 805 810 815	2448
TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr 820 825 830	2496
CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys 835 840 845	2544
CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu 850 855 860	2592
CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu 865 870 875 880	2640
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly 885 890 895	2688
ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr 900 905 910	2736
AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn	2784

915	920	925	
GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC			2832
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser			
930	935	940	
GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC			2880
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly			
945	950	955	960
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG			2928
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu			
	965	970	975
AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC			2976
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe			
	980	985	990
GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG TAC AAG TAA			3024
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys			
995	1000	1005	

(2) INFORMATION FOR SEQ ID NO:153:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1007 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

Met Ser Trp Ser Pro Ser Leu Thr Thr Gln Thr Cys Gly Ala Trp Glu			
1	5	10	15
Met Lys Glu Arg Leu Gly Thr Gly Gly Phe Gly Asn Val Ile Arg Trp			
	20	25	30
His Asn Gln Glu Thr Gly Glu Gln Ile Ala Ile Lys Gln Cys Arg Gln			
	35	40	45
Glu Leu Ser Pro Arg Asn Arg Glu Arg Trp Cys Leu Glu Ile Gln Ile			
	50	55	60
Met Arg Arg Leu Thr His Pro Asn Val Val Ala Ala Arg Asp Val Pro			
	65	70	75
Glu Gly Met Gln Asn Leu Ala Pro Asn Asp Leu Pro Leu Leu Ala Met			
	85	90	95
Glu Tyr Cys Gln Gly Gly Asp Leu Arg Lys Tyr Leu Asn Gln Phe Glu			
	100	105	110
Asn Cys Cys Gly Leu Arg Glu Gly Ala Ile Leu Thr Leu Leu Ser Asp			
	115	120	125
Ile Ala Ser Ala Leu Arg Tyr Leu His Glu Asn Arg Ile Ile His Arg			
	130	135	140

Asp Leu Lys Pro Glu Asn Ile Val Leu Gln Gln Gly Glu Gln Arg Leu
 145 150 155 160
 Ile His Lys Ile Ile Asp Leu Gly Tyr Ala Lys Glu Leu Asp Gln Gly
 165 170 175
 Ser Leu Cys Thr Ser Phe Val Gly Thr Leu Gln Tyr Leu Ala Pro Glu
 180 185 190
 Leu Leu Glu Gln Gln Lys Tyr Thr Val Thr Val Asp Tyr Trp Ser Phe
 195 200 205
 Gly Thr Leu Ala Phe Glu Cys Ile Thr Gly Phe Arg Pro Phe Leu Pro
 210 215 220
 Asn Trp Gln Pro Val Gln Trp His Ser Lys Val Arg Gln Lys Ser Glu
 225 230 235 240
 Val Asp Ile Val Val Ser Glu Asp Leu Asn Gly Thr Val Lys Phe Ser
 245 250 255
 Ser Ser Leu Pro Tyr Pro Asn Asn Leu Asn Ser Val Leu Ala Glu Arg
 260 265 270
 Leu Glu Lys Trp Leu Gln Leu Met Leu Met Trp His Pro Arg Gln Arg
 275 280 285
 Gly Thr Asp Pro Thr Tyr Gly Pro Asn Gly Cys Phe Lys Ala Leu Asp
 290 295 300
 Asp Ile Leu Asn Leu Lys Leu Val His Ile Leu Asn Met Val Thr Gly
 305 310 315 320
 Thr Ile His Thr Tyr Pro Val Thr Glu Asp Glu Ser Leu Gln Ser Leu
 325 330 335
 Lys Ala Arg Ile Gln Gln Asp Thr Gly Ile Pro Glu Glu Asp Gln Glu
 340 345 350
 Leu Leu Gln Glu Ala Gly Leu Ala Leu Ile Pro Asp Lys Pro Ala Thr
 355 360 365
 Gln Cys Ile Ser Asp Gly Lys Leu Asn Glu Gly His Thr Leu Asp Met
 370 375 380
 Asp Leu Val Phe Leu Phe Asp Asn Ser Lys Ile Thr Tyr Glu Thr Gln
 385 390 395 400
 Ile Ser Pro Arg Pro Gln Pro Glu Ser Val Ser Cys Ile Leu Gln Glu
 405 410 415
 Pro Lys Arg Asn Leu Ala Phe Phe Gln Leu Arg Lys Val Trp Gly Gln
 420 425 430
 Val Trp His Ser Ile Gln Thr Leu Lys Glu Asp Cys Asn Arg Leu Gln
 435 440 445
 Gln Gly Gln Arg Ala Ala Met Met Asn Leu Leu Arg Asn Asn Ser Cys
 450 455 460
 Leu Ser Lys Met Lys Asn Ser Met Ala Ser Met Ser Gln Gln Leu Lys
 465 470 475 480
 Ala Lys Leu Asp Phe Phe Lys Thr Ser Ile Gln Ile Asp Leu Glu Lys
 485 490 495
 Tyr Ser Glu Gln Thr Glu Phe Gly Ile Thr Ser Asp Lys Leu Leu Leu
 500 505 510
 Ala Trp Arg Glu Met Glu Gln Ala Val Glu Leu Cys Gly Arg Glu Asn
 515 520 525
 Glu Val Lys Leu Leu Val Glu Arg Met Met Ala Leu Gln Thr Asp Ile
 530 535 540
 Val Asp Leu Gln Arg Ser Pro Met Gly Arg Lys Gln Gly Gly Thr Leu
 545 550 555 560
 Asp Asp Leu Glu Glu Gln Ala Arg Glu Leu Tyr Arg Arg Leu Arg Glu
 565 570 575

Lys Pro Arg Asp Gln Arg Thr Glu Gly Asp Ser Gln Glu Met Val Arg
 580 585 590
 Leu Leu Leu Gln Ala Ile Gln Ser Phe Glu Lys Lys Val Arg Val Ile
 595 600 605
 Tyr Thr Gln Leu Ser Lys Thr Val Val Cys Lys Gln Lys Ala Leu Glu
 610 615 620
 Leu Leu Pro Lys Val Glu Glu Val Val Ser Leu Met Asn Glu Asp Glu
 625 630 635 640
 Lys Thr Val Val Arg Leu Gln Glu Lys Arg Gln Lys Glu Leu Trp Asn
 645 650 655
 Leu Leu Lys Ile Ala Cys Ser Lys Val Arg Gly Pro Val Ser Gly Ser
 660 665 670
 Pro Asp Ser Met Asn Ala Ser Arg Leu Ser Gln Pro Gly Gln Leu Met
 675 680 685
 Ser Gln Pro Ser Thr Ala Ser Asn Ser Leu Pro Glu Pro Ala Lys Lys
 690 695 700
 Ser Glu Glu Leu Val Ala Glu Ala His Asn Leu Cys Thr Leu Leu Glu
 705 710 715 720
 Asn Ala Ile Gln Asp Thr Val Arg Glu Gln Asp Gln Ser Phe Thr Ala
 725 730 735
 Leu Asp Trp Ser Trp Leu Gln Thr Glu Glu Glu Glu His Ser Cys Leu
 740 745 750
 Glu Gln Ala Ser Trp Val Pro Arg Ala Arg Asp Pro Pro Val Ala Thr
 755 760 765
 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
 770 775 780
 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
 785 790 795 800
 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
 805 810 815
 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
 820 825 830
 Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
 835 840 845
 Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
 850 855 860
 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
 865 870 875 880
 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
 885 890 895
 Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
 900 905 910
 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
 915 920 925
 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
 930 935 940
 Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
 945 950 955 960
 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
 965 970 975
 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
 980 985 990
 Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
 995 1000 1005

(2) INFORMATION FOR SEQ ID NO:154:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2793 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...2790
- (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

ATG ATG CAC GTG AAT AAT TTT CCC TTT AGA AGG CAT TCC TGG ATA TGT	48
Met Met His Val Asn Asn Phe Pro Phe Arg Arg His Ser Trp Ile Cys	
1 5 10 15	
TTT GAT GTG GAC AAT GGC ACA TCT GCG GGA CGG AGT CCC TTG GAT CCC	96
Phe Asp Val Asp Asn Gly Thr Ser Ala Gly Arg Ser Pro Leu Asp Pro	
20 25 30	
ATG ACC AGC CCA GGA TCC GGG CTA ATT CTC CAA GCA AAT TTT GTC CAC	144
Met Thr Ser Pro Gly Ser Gly Leu Ile Leu Gln Ala Asn Phe Val His	
35 40 45	
AGT CAA CGA CGG GAG TCC TTC CTG TAT CGA TCC GAC AGC GAT TAT GAC	192
Ser Gln Arg Arg Glu Ser Phe Leu Tyr Arg Ser Asp Ser Asp Tyr Asp	
50 55 60	
CTC TCT CCA AAG TCT ATG TCC CGG AAC TCC TCC ATT GCC AGT GAT ATA	240
Leu Ser Pro Lys Ser Met Ser Arg Asn Ser Ser Ile Ala Ser Asp Ile	
65 70 75 80	
CAC GGA GAT GAC TTG ATT GTG ACT CCA TTT GCT CAG GTC TTG GCC AGT	288
His Gly Asp Asp Leu Ile Val Thr Pro Phe Ala Gln Val Leu Ala Ser	
85 90 95	
CTG CGA ACT GTA CGA AAC AAC TTT GCT GCA TTA ACT AAT TTG CAA GAT	336
Leu Arg Thr Val Arg Asn Asn Phe Ala Ala Leu Thr Asn Leu Gln Asp	
100 105 110	
CGA GCA CCT AGC AAA AGA TCA CCC ATG TGC AAC CAA CCA TCC ATC AAC	384
Arg Ala Pro Ser Lys Arg Ser Pro Met Cys Asn Gln Pro Ser Ile Asn	
115 120 125	
AAA GCC ACC ATA ACA GAG GAG GCC TAC CAG AAA CTG GCC AGC GAG ACC	432
Lys Ala Thr Ile Thr Glu Glu Ala Tyr Gln Lys Leu Ala Ser Glu Thr	
130 135 140	

CTG GAG GAG CTG GAC TGG TGT CTG GAC CAG CTA GAG ACC CTA CAG ACC	480
Leu Glu Glu Leu Asp Trp Cys Leu Asp Gln Leu Glu Thr Leu Gln Thr	
145 150 155 160	
AGG CAC TCC GTC AGT GAG ATG GCC TCC AAC AAG TTT AAA AGG ATG CTT	528
Arg His Ser Val Ser Glu Met Ala Ser Asn Lys Phe Lys Arg Met Leu	
165 170 175	
AAT CGG GAG CTC ACC CAT CTC TCT GAA ATG AGT CGG TCT GGA AAT CAA	576
Asn Arg Glu Leu Thr His Leu Ser Glu Met Ser Arg Ser Gly Asn Gln	
180 185 190	
GTG TCA GAG TTT ATA TCA AAC ACA TTC TTA GAT AAG CAA CAT GAA GTG	624
Val Ser Glu Phe Ile Ser Asn Thr Phe Leu Asp Lys Gln His Glu Val	
195 200 205	
GAA ATT CCT TCT CCA ACT CAG AAG GAA AAG GAG AAA AAG AAA AGA CCA	672
Glu Ile Pro Ser Pro Thr Gln Lys Glu Lys Glu Lys Lys Lys Arg Pro	
210 215 220	
ATG TCT CAG ATC AGT GGA GTC AAG AAA TTG ATG CAC AGC TCT AGT CTG	720
Met Ser Gln Ile Ser Gly Val Lys Lys Leu Met His Ser Ser Ser Leu	
225 230 235 240	
ACT AAT TCA AGT ATC CCA AGG TTT GGA GTT AAA ACT GAA CAA GAA GAT	768
Thr Asn Ser Ser Ile Pro Arg Phe Gly Val Lys Thr Glu Gln Glu Asp	
245 250 255	
GTC CTT GCC AAG GAA CTA GAA GAT GTG AAC AAA TGG GGT CTT CAT GTT	816
Val Leu Ala Lys Glu Leu Glu Asp Val Asn Lys Trp Gly Leu His Val	
260 265 270	
TTC AGA ATA GCA GAG TTG TCT GGT AAC CGG CCC TTG ACT GTT ATC ATG	864
Phe Arg Ile Ala Glu Leu Ser Gly Asn Arg Pro Leu Thr Val Ile Met	
275 280 285	
CAC ACC ATT TTT CAG GAA CGG GAT TTA TTA AAA ACA TTT AAA ATT CCA	912
His Thr Ile Phe Gln Glu Arg Asp Leu Leu Lys Thr Phe Lys Ile Pro	
290 295 300	
GTA GAT ACT TTA ATT ACA TAT CTT ATG ACT CTC GAA GAC CAT TAC CAT	960
Val Asp Thr Leu Ile Thr Tyr Leu Met Thr Leu Glu Asp His Tyr His	
305 310 315 320	
GCT GAT GTG GCC TAT CAC AAC AAT ATC CAT GCT GCA GAT GTT GTC CAG	1008
Ala Asp Val Ala Tyr His Asn Asn Ile His Ala Ala Asp Val Val Gln	
325 330 335	
TCT ACT CAT GTG CTA TTA TCT ACA CCT GCT TTG GAG GCT GTG TTT ACA	1056
Ser Thr His Val Leu Leu Ser Thr Pro Ala Leu Glu Ala Val Phe Thr	
340 345 350	
GAT TTG GAG ATT CTT GCA GCA ATT TTT GCC AGT GCA ATA CAT GAT GTA	1104
Asp Leu Glu Ile Leu Ala Ala Ile Phe Ala Ser Ala Ile His Asp Val	

355	360	365	
GAT CAT CCT GGT GTG TCC AAT CAA TTT CTG ATC AAT ACA AAC TCT GAA			1152
Asp His Pro Gly Val Ser Asn Gln Phe Leu Ile Asn Thr Asn Ser Glu			
370	375	380	
CTT GCC TTG ATG TAC AAT GAT TCC TCA GTC TTA GAG AAC CAT CAT TTG			1200
Leu Ala Leu Met Tyr Asn Asp Ser Ser Val Leu Glu Asn His His Leu			
385	390	395	400
GCT GTG GGC TTT AAA TTG CTT CAG GAA GAA AAC TGT GAC ATT TTC CAG			1248
Ala Val Gly Phe Lys Leu Leu Gln Glu Glu Asn Cys Asp Ile Phe Gln			
405	410	415	
AAT TTG ACC AAA AAA CAA AGA CAA TCT TTA AGG AAA ATG GTC ATT GAC			1296
Asn Leu Thr Lys Lys Gln Arg Gln Ser Leu Arg Lys Met Val Ile Asp			
420	425	430	
ATC GTA CTT GCA ACA GAT ATG TCA AAA CAC ATG AAT CTA CTG GCT GAT			1344
Ile Val Leu Ala Thr Asp Met Ser Lys His Met Asn Leu Leu Ala Asp			
435	440	445	
TTG AAG ACT ATG GTT GAA ACT AAG AAA GTG ACA AGC TCT GGA GTT CTT			1392
Leu Lys Thr Met Val Glu Thr Lys Lys Val Thr Ser Ser Gly Val Leu			
450	455	460	
CTT CTT GAT AAT TAT TCC GAT AGG ATT CAG GTT CTT CAG AAT ATG GTG			1440
Leu Leu Asp Asn Tyr Ser Asp Arg Ile Gln Val Leu Gln Asn Met Val			
465	470	475	480
CAC TGT GCA GAT CTG AGC AAC CCA ACA AAG CCT CTC CAG CTG TAC CGC			1488
His Cys Ala Asp Leu Ser Asn Pro Thr Lys Pro Leu Gln Leu Tyr Arg			
485	490	495	
CAG TGG ACG GAC CGG ATA ATG GAG GAG TTC TTC CGC CAA GGA GAC CGA			1536
Gln Trp Thr Asp Arg Ile Met Glu Glu Phe Phe Arg Gln Gly Asp Arg			
500	505	510	
GAG AGG GAA CGT GGC ATG GAG ATA AGC CCC ATG TGT GAC AAG CAC AAT			1584
Glu Arg Glu Arg Gly Met Glu Ile Ser Pro Met Cys Asp Lys His Asn			
515	520	525	
GCT TCC GTG GAA AAA TCA CAG GTG GGC TTC ATA GAC TAT ATT GTT CAT			1632
Ala Ser Val Glu Lys Ser Gln Val Gly Phe Ile Asp Tyr Ile Val His			
530	535	540	
CCC CTC TGG GAG ACA TGG GCA GAC CTC GTC CAC CCT GAC GCC CAG GAT			1680
Pro Leu Trp Glu Thr Trp Ala Asp Leu Val His Pro Asp Ala Gln Asp			
545	550	555	560
ATT TTG GAC ACT TTG GAG GAC AAT CGT GAA TGG TAC CAG AGC ACA ATC			1728
Ile Leu Asp Thr Leu Glu Asp Asn Arg Glu Trp Tyr Gln Ser Thr Ile			
565	570	575	

CCT CAG AGC CCC TCT CCT GCA CCT GAT GAC CCA GAG GAG GGC CGG CAG Pro Gln Ser Pro Ser Pro Ala Pro Asp Asp Pro Glu Glu Gly Arg Gln 580 585 590	1776
GGT CAA ACT GAG AAA TTC CAG TTT GAA CTA ACT TTA GAG GAA GAT GGT Gly Gln Thr Glu Lys Phe Gln Phe Glu Leu Thr Leu Glu Glu Asp Gly 595 600 605	1824
GAG TCA GAC ACG GAA AAG GAC AGT GGC AGT CAA GTG GAA GAA GAC ACT Glu Ser Asp Thr Glu Lys Asp Ser Gly Ser Gln Val Glu Glu Asp Thr 610 615 620	1872
AGC TGC AGT GAC TCC AAG ACT CTT TGT ACT CAA GAC TCA GAG TCT ACT Ser Cys Ser Asp Ser Lys Thr Leu Cys Thr Gln Asp Ser Glu Ser Thr 625 630 635 640	1920
GAA ATT CCC CTT GAT GAA CAG GTT GAA GAG GAG GCA GTA GGG GAA GAA Glu Ile Pro Leu Asp Glu Gln Val Glu Glu Glu Ala Val Gly Glu Glu 645 650 655	1968
GAG GAA AGC CAG CCT GAA GCC TGT GTC ATA GAT GAT CGT TCT CCT GAC Glu Glu Ser Gln Pro Glu Ala Cys Val Ile Asp Asp Arg Ser Pro Asp 660 665 670	2016
ACG ACG GGA ATT CTG CAG TCG ACG GTA CCG CGG GCC CGG GAT CCA CCG Thr Thr Gly Ile Leu Gln Ser Thr Val Pro Arg Ala Arg Asp Pro Pro 675 680 685	2064
GTC GCC ACC ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG GTG GTG Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val 690 695 700	2112
CCC ATC CTG GTC GAG CTG GAC GGC GAC GTA AAC GGC CAC AAG TTC AGC Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser 705 710 715 720	2160
GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu 725 730 735	2208
AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu 740 745 750	2256
GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp 755 760 765	2304
CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr 770 775 780	2352
GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr	2400

785	790	795	800	
CGC GCC GAG GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG				2448
Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu				
805		810	815	
CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG				2496
Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys				
820		825	830	
CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG				2544
Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys				
835		840	845	
CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG				2592
Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu				
850		855	860	
GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC				2640
Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile				
865		870	875	880
GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG				2688
Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln				
885		890	895	
TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG				2736
Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu				
900		905	910	
CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CTC GGC ATG GAC GAG CTG				2784
Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu				
915		920	925	
TAC AAG TAA				2793
Tyr Lys				
930				

(2) INFORMATION FOR SEQ ID NO:155:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 930 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

Met	Met	His	Val	Asn	Asn	Phe	Pro	Phe	Arg	Arg	His	Ser	Trp	Ile	Cys
1					5				10					15	

Phe	Asp	Val	Asp	Asn	Gly	Thr	Ser	Ala	Gly	Arg	Ser	Pro	Leu	Asp	Pro	20	25	30
Met	Thr	Ser	Pro	Gly	Ser	Gly	Leu	Ile	Leu	Gln	Ala	Asn	Phe	Val	His	35	40	45
Ser	Gln	Arg	Arg	Glu	Ser	Phe	Leu	Tyr	Arg	Ser	Asp	Ser	Asp	Tyr	Asp	50	55	60
Leu	Ser	Pro	Lys	Ser	Met	Ser	Arg	Asn	Ser	Ser	Ile	Ala	Ser	Asp	Ile	65	70	75
His	Gly	Asp	Asp	Leu	Ile	Val	Thr	Pro	Phe	Ala	Gln	Val	Leu	Ala	Ser	85	90	95
Leu	Arg	Thr	Val	Arg	Asn	Asn	Phe	Ala	Ala	Leu	Thr	Asn	Leu	Gln	Asp	100	105	110
Arg	Ala	Pro	Ser	Lys	Arg	Ser	Pro	Met	Cys	Asn	Gln	Pro	Ser	Ile	Asn	115	120	125
Lys	Ala	Thr	Ile	Thr	Glu	Glu	Ala	Tyr	Gln	Lys	Leu	Ala	Ser	Glu	Thr	130	135	140
Leu	Glu	Glu	Leu	Asp	Trp	Cys	Leu	Asp	Gln	Leu	Glu	Thr	Leu	Gln	Thr	145	150	155
Arg	His	Ser	Val	Ser	Glu	Met	Ala	Ser	Asn	Lys	Phe	Lys	Arg	Met	Leu	165	170	175
Asn	Arg	Glu	Leu	Thr	His	Leu	Ser	Glu	Met	Ser	Arg	Ser	Gly	Asn	Gln	180	185	190
Val	Ser	Glu	Phe	Ile	Ser	Asn	Thr	Phe	Leu	Asp	Lys	Gln	His	Glu	Val	195	200	205
Glu	Ile	Pro	Ser	Pro	Thr	Gln	Lys	Glu	Lys	Glu	Lys	Lys	Lys	Arg	Pro	210	215	220
Met	Ser	Gln	Ile	Ser	Gly	Val	Lys	Lys	Leu	Met	His	Ser	Ser	Ser	Leu	225	230	235
Thr	Asn	Ser	Ser	Ile	Pro	Arg	Phe	Gly	Val	Lys	Thr	Glu	Gln	Glu	Asp	245	250	255
Val	Leu	Ala	Lys	Glu	Leu	Glu	Asp	Val	Asn	Lys	Trp	Gly	Leu	His	Val	260	265	270
Phe	Arg	Ile	Ala	Glu	Leu	Ser	Gly	Asn	Arg	Pro	Leu	Thr	Val	Ile	Met	275	280	285
His	Thr	Ile	Phe	Gln	Glu	Arg	Asp	Leu	Leu	Lys	Thr	Phe	Lys	Ile	Pro	290	295	300
Val	Asp	Thr	Leu	Ile	Thr	Tyr	Leu	Met	Thr	Leu	Glu	Asp	His	Tyr	His	305	310	315
Ala	Asp	Val	Ala	Tyr	His	Asn	Asn	Ile	His	Ala	Ala	Asp	Val	Val	Gln	325	330	335
Ser	Thr	His	Val	Leu	Leu	Ser	Thr	Pro	Ala	Leu	Glu	Ala	Val	Phe	Thr	340	345	350
Asp	Leu	Glu	Ile	Leu	Ala	Ala	Ile	Phe	Ala	Ser	Ala	Ile	His	Asp	Val	355	360	365
Asp	His	Pro	Gly	Val	Ser	Asn	Gln	Phe	Leu	Ile	Asn	Thr	Asn	Ser	Glu	370	375	380
Leu	Ala	Leu	Met	Tyr	Asn	Asp	Ser	Ser	Val	Leu	Glu	Asn	His	His	Leu	385	390	395
Ala	Val	Gly	Phe	Lys	Leu	Leu	Gln	Glu	Glu	Asn	Cys	Asp	Ile	Phe	Gln	405	410	415
Asn	Leu	Thr	Lys	Lys	Gln	Arg	Gln	Ser	Leu	Arg	Lys	Met	Val	Ile	Asp	420	425	430
Ile	Val	Leu	Ala	Thr	Asp	Met	Ser	Lys	His	Met	Asn	Leu	Leu	Ala	Asp	435	440	445

Leu Lys Thr Met Val Glu Thr Lys Lys Val Thr Ser Ser Gly Val Leu
 450 455 460
 Leu Leu Asp Asn Tyr Ser Asp Arg Ile Gln Val Leu Gln Asn Met Val
 465 470 475 480
 His Cys Ala Asp Leu Ser Asn Pro Thr Lys Pro Leu Gln Leu Tyr Arg
 485 490 495
 Gln Trp Thr Asp Arg Ile Met Glu Glu Phe Phe Arg Gln Gly Asp Arg
 500 505 510
 Glu Arg Glu Arg Gly Met Glu Ile Ser Pro Met Cys Asp Lys His Asn
 515 520 525
 Ala Ser Val Glu Lys Ser Gln Val Gly Phe Ile Asp Tyr Ile Val His
 530 535 540
 Pro Leu Trp Glu Thr Trp Ala Asp Leu Val His Pro Asp Ala Gln Asp
 545 550 555 560
 Ile Leu Asp Thr Leu Glu Asp Asn Arg Glu Trp Tyr Gln Ser Thr Ile
 565 570 575
 Pro Gln Ser Pro Ser Pro Ala Pro Asp Asp Pro Glu Glu Gly Arg Gln
 580 585 590
 Gly Gln Thr Glu Lys Phe Gln Phe Glu Leu Thr Leu Glu Glu Asp Gly
 595 600 605
 Glu Ser Asp Thr Glu Lys Asp Ser Gly Ser Gln Val Glu Glu Asp Thr
 610 615 620
 Ser Cys Ser Asp Ser Lys Thr Leu Cys Thr Gln Asp Ser Glu Ser Thr
 625 630 635 640
 Glu Ile Pro Leu Asp Glu Gln Val Glu Glu Glu Ala Val Gly Glu Glu
 645 650 655
 Glu Glu Ser Gln Pro Glu Ala Cys Val Ile Asp Asp Arg Ser Pro Asp
 660 665 670
 Thr Thr Gly Ile Leu Gln Ser Thr Val Pro Arg Ala Arg Asp Pro Pro
 675 680 685
 Val Ala Thr Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val
 690 695 700
 Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser
 705 710 715 720
 Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu
 725 730 735
 Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu
 740 745 750
 Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp
 755 760 765
 His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr
 770 775 780
 Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr
 785 790 795 800
 Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu
 805 810 815
 Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys
 820 825 830
 Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys
 835 840 845
 Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu
 850 855 860
 Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile
 865 870 875 880

Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln
 885 890 895
 Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu
 900 905 910
 Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu
 915 920 925
 Tyr Lys
 930

(2) INFORMATION FOR SEQ ID NO:156:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

GTAAGCTTCG AACATGATGC ACGTGAATAA TTTTCCC

37

(2) INFORMATION FOR SEQ ID NO:157:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

GTAAGCTTCG AACATGGAGG CAGAGGGCAG CAGC

34

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

GTAAGCTTCG AACATGGCTC AGCAGACAAG CCCG

34

(2) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

GTGAATTCCC GTCGTGTCAG GAGAAGCATC ATCTATG

37

(2) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

GTGAATTCAA CCATGGAGCG GGCC

24

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

GTGGTACCCA GTTCGCTTG GCC

23

(2) INFORMATION FOR SEQ ID NO:162:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

GTCTCGAGGC AAGATGGCTG ACCC

24

(2) INFORMATION FOR SEQ ID NO:163:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

GTGGATCCGA GCTCTTGACT TCGGG

25

(2) INFORMATION FOR SEQ ID NO:164:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

GTAAGCTTAC ATGAGCTGGT CACCTTCCT G

31

(2) INFORMATION FOR SEQ ID NO:165:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

GTGGTACCCA TGAGGCCTGC TCCAG

25

METHOD AND APPARATUS FOR HIGH DENSITY FORMAT SCREENING FOR BIOACTIVE MOLECULES

FIELD OF THE INVENTION

The invention relates to a method and apparatus for screening large numbers of molecules for biological activities.

BACKGROUND OF THE INVENTION

Current technology is able to generate large numbers of molecules which may possess potential therapeutic value. Compounds having potentially interesting biological activity may be products of combinatorial or traditional chemistry, a natural product, proteins isolated by one- or two-dimensional gel electrophoresis, or compounds secreted from or expressed by natural or genetically modified animal, plant, microbial or fungal cells (or parts thereof), or displayed by natural or genetically modified viral or phage particles.

Screening methods have been developed which achieve very high throughputs of test compounds. Such methods are termed Ultra High Throughput Screening (UHTS). The present generation of UHTS machines rely upon essentially serial additions of test compounds, usually one test compound per discrete test well. Test well array densities range from between 96 to 3456 wells per plate. Such UHTS machines require sophisticated technologies to dispense microvolumes of many different fluids to selected locations, and also require that the detecting surface for each test molecule generally be separated from other detecting surfaces within the array.

There is a need to develop a method which allows simultaneous screening of large numbers of test compounds for biological activity and potential therapeutic use while avoiding the complications associated with dispensing multiple fluid microvolumes.

BRIEF SUMMARY OF THE INVENTION

The invention is directed to a screening method which eliminates the need for delivering microfluid volumes and allows simultaneous parallel screening of large numbers of test compounds. Accordingly, the invention is drawn to a method for screening test

compounds for bioactivity, by contacting an array of test compounds with a detector layer capable of detecting bioactivity, wherein a cell response is indicative of bioactivity.

The method of the invention is a high throughput system for parallel screening of a large number of test compounds. In one embodiment of the method of the invention, 96 to 10,000 test compounds are simultaneously screened for bioactivity in an assay; in a more specific embodiment, 96 to 3456 test compounds are simultaneously screened for bioactivity.

In a more specific embodiment, invention is drawn to a method for screening test compounds for bioactivity, comprising:

(a) contacting a solid support comprising an array of test compounds with a liquid layer, wherein the liquid layer is in immediate contact with a detector layer and wherein each test compound comes into contact with a localized portion of the liquid layer; and

(b) registering a response of the detector layer to the test compound, wherein a bioactive test compound is identified.

By "high throughput screening" is meant a method able to screen large number of test compounds for biological activity within a given machine time (i.e. at a rate anywhere from 100 to 100,000 compounds per hour per machine).

The term "parallel screening" refers to a method by which very many compounds are applied simultaneously to the detector layer, and similarly, signals from that detector layer are collected contemporaneously rather than sequentially.

By "array" is meant a regular two-dimensional arrangement of test compounds by which compounds are disposed at the nodes of a rectilinear grid pattern whereby a compound position can be identified by a simple 2-dimensional coordinate.

A "detector layer" means any two-dimensional system which can be used to report biologically relevant information. In one specific embodiment of the method of the invention the detector layer is a monolayer of living cells loaded with a fluorescent reporter dye such as Fluo-3.

By "bioactive" or "bioactivity" is meant an action or influence of a test compound upon the detector layer which results in a response from the detector layer that has direct biological significance or can be interpreted as being a biologically relevant response. Bioactive agents have the ability to effect physiological parameters of living cells and tissues. Bioactivity includes inducing or suppressing the expression of a protein, activating or inhibiting transcription of a gene, and/or effecting cellular function(s) such as, for example,

intracellular movement and storage of calcium ions, and membrane transportation.

The capacity of a test compound to affect a detector layer, i.e. bioactivity, may be determined in a number of ways known to the art. In specific embodiments of the method of the invention, bioactivity is determined by changes or movements of fluorescent probes present in the detector layer which indicate changes in ionic content, cell metabolism, growth or viability. In a preferred method of the invention, living cells form the detector layer and have specific protein components tagged with a fluorescent agent, such as green fluorescent protein (GFP); changes in GFP fluorescence or distribution within cells indicate a particular cellular response which may be selected for identification of bioactivity.

The phrase "a change in fluorescence" means any change in absorption properties, such as wavelength and intensity, or any change in spectral properties of the emitted light, such as a change of wavelength, fluorescence lifetime, intensity or polarization.

A "solid support comprising an array of multiple test compounds" or similar terms, mean a fixed matrix to which test compounds have been fixed. As an example, the solid support of the invention includes a membrane or other surface comprising an array of printed test compounds. In one specific embodiment of the invention, the test compounds are deposited as discrete spots on a porous track-etched polycarbonate membrane 10 to 20 microns thickness, the spots being between 10 microns to 2 mm diameter. The quantity of compound contained in each discrete spot will depend on the concentration of the stock solution from which it was derived, and the volume of that stock solution applied to the support. In another specific embodiment of the invention, compounds are printed onto a non-porous solid support which is optically clear.

By "test compounds" is meant a fixed array of compounds to be screened for ability to effect physiological parameters of a cell or tissue. In one embodiment, the test compounds are proteins or peptides generated by combinatorial protein chemical methods known to the art. In another embodiment, the test compounds are chemical compounds generated by combinatorial chemistry methods known in the art. In another embodiments, the test compounds are chemical compounds which are naturally occurring compounds more or less purified from their native state, are the products of genetically engineered cells, or are viral or bacteriophage particles engineered to display compounds upon their surfaces (phage display).

In one embodiment, the detector layer is an undemarcated area of living cells growing on a flat culture surface. The cells on this surface may or may not be grown to confluence,

may be transformed and/or engineered cells, or directly derived from animal tissues and grown as primary cell culture.

In one embodiment, a test compound reaches the detector layer by diffusion through a porous membrane to a liquid layer immediately overlaying the detector layer. A variety of commercially available porous membranes are useful in the method of the invention. A preferred porous membrane is a track-etched polyester or polycarbonate support in which parallel channels of identical size are formed by a selective etching process following exposure of the membrane to a source of high energy ions. The method of the invention allows each test compound affixed to a solid support to come into contact with a limited fluid volume, which fluid volume is in immediate contact with the detector layer. In one embodiment, each test compound contacts the detector layer by diffusion through a liquid-containing channel directly adjacent to the detector layer.

One advantage of the method of the invention is that it allows massive parallel screening of a large array of test compounds for biological activity. When living cells are the detector layer of the invention, they are maintained under physiologically viable conditions. Provision of these conditions requires the use of solutions able to supply essential nutrients and buffer pH changes normal to the continued growth of living cells. Such solutions may be complete cell culture media (i.e. any of those commercially available, for instance from Life Technologies Ltd.), optionally supplemented with antibiotics and serum preparations for optimal cell growth conditions. Buffer solutions may also be of the type known as "chemically defined". Cells will also require controlled temperature conditions, in the range 20° to 37°C, and the provision of gases essential to continued cell growth and maintenance of buffer capacity (O₂, and optionally 5% CO₂, depending on the type of buffer being used).

These and other objectives, advantages, and features of the invention will become apparent to those persons skilled in the art upon reading the details of the method as more fully described below.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing features of the present invention may be fully understood from the following detailed disclosure of a specific preferred embodiment in conjunction with the accompanying drawings in which:

Fig. 1 is a schematic representation of the apparatus useful in one specific embodiment of the invention: Light from a high energy light source **1** is collected and collimated by unit **2**, directed through a shutter assembly **3** and passes through a excitation filter-changer **4**. A light guide **5** directs excitation light into the lensing and epi-illumination optics housed in unit **7**. Excitation light emerging from **7** illuminates the horizontal detector layer located in the multi-component assembly having two solid layers **10** and **11** fixed relative to a supporting stage unit **8**. Layer **11** is moved vertically downward on guide pins (17 Fig. 2b) controlled by arm **12** driven by unit **13**. Four sprung contacts **14** attached to **12** press upon the frame of layer **11** to drive it downwards as arm **12** descends. Specified devices (**3, 4, 9, 13, 15, 16**) are controlled by central processing unit **6** which issues commands and collects data and status information from the devices attached to it. Unit **6** includes a central processing unit, RAM, multi-channel serial input/output cards with onboard A/D and D/A converters, one of which cards controls the camera **16** and captures images from it.

Figs. 2a-c: Figs. 2a and 2b are side view of the test stage (not to scale); Fig. 2c is a top view of the test stage. A supporting stage **8** has a rectangular central aperture the shape and size of which is the same as the area **19** of Fig. 2c. The position of stage **8** is adjusted in the horizontal and vertical axes by the 3-axis positioner **9**. Components of the test stage shown include, solution layer **18**, (not shown: detector layer **20** and array of test compounds **21** in Figs 3 and 4). The array **21** is held away from the liquid layer by pins **17** which pass through holes (**24** in Fig. 5) in the corners of the frame **11**. Arm **12** is moved down by the drive unit **13**, and the four sprung contacts **14** it bears exert pressure on the frame **11** moving it down the guide pins **17** and into close proximity to the upper surface of **10**, from which it is separated by a thin liquid layer **18**.

Fig. 3 is a schematic showing the relative positions of the different layers in the test-array/detector layers used in one specific embodiment. The layers are depicted in apposition, as they would appear after arm **12** has pushed component **11** down the support pins **17**. An array of discrete spots of test compounds **21** on a porous membrane **19** is in contact with a liquid layer **18** overlaying the detector layer **20** which is supported by an optically transparent

solid substrate **10**. The compounds fill the parallel capillary spaces in the track-etched membrane **22**.

Fig. 4 is a schematic drawing of a second embodiment of the screening method of the invention. The layers are depicted in apposition, as they would appear after arm **12** has pushed component **11** down the support pins **17**. A detector layer **20** supported on an optically clear porous membrane **19**, and overlaid by a liquid layer **23**, is placed onto an optically clear solid substrate **10** bearing an array of test compounds **21**. The thin space **18** between components **19** and **10** is filled with solution from **23** which has passed through the porous membrane **19**. Bioactivity is detected by measuring changes in fluorescence of the detector layer resulting from responses to the diffusion of test compounds through the porous membrane to the detector layer.

Figs. 5a-c are schematics illustrating transfer printing of an array of compounds onto a surface of a track-etched membrane. Compounds are stored in 16 separate 96-well microtitre plates and defined amounts are transferred simultaneously by a 96-pin printing head to the surface **19** (Fig. 5a). The contents of each successive 96-well plate are printed at a slightly offset position, generating an array after 4 such printing operations (Fig. 5b), and a full array of 1536 compounds after 16 printing operations (Fig. 5c).

DETAILED DESCRIPTION

Before the present method and solutions used in the method are described, it is to be understood that this invention is not limited to particular methods, components, or solutions described, as such methods, components, and solutions may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and

materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited.

Generally, the invention is drawn to a method for high throughput screening of test compounds, by contacting a solid support comprising an array of multiple test compounds with a detector layer, wherein each test compound comes into contact with a localized liquid which is in contact with a detector layer, and detecting a response of the detector layer to the test compound, wherein a bioactive test compound is identified.

The high density format screening system (HDFS) of the invention, rests in part on the realization that the delivery of test compounds to detector surfaces can be greatly simplified by doing away with the need for complicated microfluidics. Test compounds are applied to the detector surface in a massively parallel manner, and the method is applicable to a large range of different types of test compounds.

Central to the specific embodiments of the method and apparatus of the invention, described below, is the use of living cells as detectors, their responses being signalled via changes in the fluorescent or luminescent properties of various specific probes located within. However many different types of detector systems could be used in place of cells in such a system, for example, appropriate variants of Scintillation Proximity Assay (SPA) systems (Amersham Pharmacia Biotech) and enzyme-linked immuno-sorbent assay (ELISA) systems (Amersham).

Test Compound Arrays

The array of test compounds is formatted to have the same dimensions as the detector surface. In one specific embodiment of the invention, array and detector layers have a width of 8 cm and length of 12.5 cm, so as to fit within the format of conventional 96-well or 384-well microtiter plates. Preparation of the test arrays will depend on their origin.

Test compounds held in formatted arrays. Current methods for the production of single compounds by combinatorial methods are under development which involve miniaturization and patterned arrays of tethered solid-phase substrates. Thus, test compounds generated by combinatorial methods can be used to synthesize an array directly or indirectly on a carrier sheet. In one embodiment, vapor phase solubilization is used to produce a test compound array on the synthetic substrate, followed by a printing process of the test

compound array on to an absorbent membrane. In this embodiment, the test array is the printed membrane. An attractive feature of this method is that multiple copies of the same test array can be produced at one time to be screened against multiple cell systems for specific activities which minimizes stock handling from library archives.

Currently most compounds to be screened come in 96-well format. However, the 96-well format can be altered by repeated off-set printings, to any chosen density of format that the transfer substrate and assay can support. The optimum density of compounds in the test array will depend very much on the fraction of compounds in an array which generate bioactive responses in the detector layer ("hit rate"). The hit rate will depend on how well the compound library being tested matches the targets in the assay. If the hit rate is low, e.g., 1:20,000 - 100,000 compounds tested, a test array with center to center spacing of 200 μm (giving 240,000 separate compounds in a 12 cm x 8 cm area) may be preferable, providing 2 to 10 hits per plate. At a spacing of 1 mm, 9,600 test compounds may be screened simultaneously.

The density of the format may be adjusted as required without requiring any changes in the hardware used to perform the re-formatting; rather, adjustment may be made in the degree of off-set and the number of print operations used per array.

Detection

Fluorescent imaging provides a way to monitor physiological responses of living cells in a non-invasive manner. Ion- and voltage-sensitive probes, as well as the new generation of recombinant fluorescent probes, for instance, hybrid proteins comprising fusions of green fluorescent protein variants (GFPs) to cellular proteins involved in intracellular signaling, can be used singly or in combination to report on many aspects of cellular microphysiology. Due to the strong fluorescence of GFP, the luminescence of cells expressing the probes may easily be detected and analyzed by employing a combination of fluorescence microscopy and image analysis. Furthermore, these probes described are easily introduced into cells, as they can be expressed in the cells of interest after transfection with a suitable expression vector.

Recombinant probes for second messengers and enzyme activity, such as kinase activity, are not only useful in basic research but also in screening programs aiming at identifying novel biologically active substances. As an example, any currently used screening program designed to find compounds that affect cAMP concentration and protein kinase

activity are based on receptor binding and/or immuno detection and/or reporter gene expression. The recombinant probes described herein, on the other hand, make it possible to develop an entirely new types of screening assays able to monitor immediate and transient changes of cAMP concentration and protein kinase activity in intact living cells.

The HDFS method of the invention monitors the response of cell populations to test compounds. Lens systems are currently available which can simultaneously epi-illuminate and image the fluorescence from areas in excess of 8.5 x 13 cm, the size of a standard 96-well plate. The detection method used herein collects a variety of fluorescent signals from all cells in a field, with responses from discrete areas of the field being apparent in the real image of the fluorescence from that field as formed on the surface of the photosensitive detector (imaging camera).

Delivery of Test Compounds to Detector Cells

In a first embodiment of the method of the invention, delivery of large arrays of test compounds to cells is achieved with test compounds which are present on or transferred to a porous carrier sheet. In specific embodiments, test compounds are printed on the carrier sheet, and the sheet is applied (overlayed) to a field of cells of the same area. The test compounds reach the detector cells by diffusion through a localized buffer layer immediately in contact with an area of the detector cell layer. This embodiment is shown in the schematic of Figs. 2 & 3.

Porous carrier sheet for delivery of test compounds: Test compound arrays are fixed onto the porous carrier sheet by a variety of methods known to the art. For example, an array of test compounds may be transferred and fixed to the carrier sheet by the method of contact printing, whereby an array of inert flat-ended pins (e.g. made of stainless steel) is used to transfer defined volumes of individual test compounds (in the range 50 nl to 2 µl) in solution form to discrete points on a dry carrier sheet.

A porous membrane useful in the delivery of test compounds is a membrane constructed of a non-absorbent material with pores of regular and defined diameter which traverse the membrane directly from the upper to the lower side. The property of orthogonal capillarity is useful in these membranes to limit lateral spread of test compounds applied to the membranes as discrete spots of liquid, since it is important that the compounds remain as discrete spots upon the membrane. A variety of membranes of different thicknesses,

materials, and pore densities are commercially available from a number of manufacturers. For example, porous membranes useful in the method of the invention include a track-etched polycarbonate or polyester membrane (Corning Costar or Whatman/Polyfiltronics). These are available in thicknesses from 6 to 23 microns, with pores of 14 to 0.015 microns, at 100,000 to 1,000,000,000 pores/cm². For delivery of test compounds with maximum ease of handling and loading of test compounds, polycarbonate membranes are preferred, particularly of a thickness of greater than 10 microns, with pores between 1 and 10 microns diameter at densities of between 20,000,000 to 100,000 pores/cm², respectively. One preferred membrane is Nucleopore® from Corning Costar.

Alternative membranes useful for the delivery of compounds include cast cellulose acetate (Membra-fil®), PTFE membranes (e.g. Filinert™), and glass fiber filters, all available from Corning Costar. These thicker membranes encourage lateral spread of liquid samples applied to their surfaces, but are thicker and could thus be used to deliver larger amounts of compounds.

Track-etched and cast cellulosic membranes may also be given hydrophilic or hydrophobic surface treatments. It is useful to have membranes whose surfaces have defined wettability properties.

When the test compound is soluble, the compound will dissolve into the buffer upon contact with the buffer medium, and directly contact the detector layer immediately underlying the buffer layer. In this embodiment, the test compounds dissolve upon contact with the buffer medium, and fall vertically onto the detector layer as a result of having a higher density than the surrounding liquor. It is generally preferred that the thin buffer layer between the test compound membrane and detector layer not be stirred significantly by convection. At the detector layer, the vertical fall of a solution of test compound is expected to spread radially by displacement and diffusion. The radial extent of a measured response may thus be used as an indicator of the bio-potency of the compounds involved.

Test compounds of limited solubility, such as those expressed on the surface of a carrier system, for instance, a cell membrane, viral or phage particle, must be brought into very close proximity, including direct contact, with the detector layers.

Buffer and Detector layer. The detector layer may be a continuous or non-continuous layer of living cells. In a specific embodiment, the detector layer is a continuous cell monolayer corresponding in size to the test compound array. In more specific embodiments,

thin glass substrate, suitably tissue culture treated is preferred for fluorescent probes requiring excitation wavelengths below 400 nm.

Living cells are maintained under physiologically viable conditions, as defined by such parameters as oxygen consumption, membrane potential, mitochondrial potential and cytoplasmic ion balance. Provision of these conditions requires the use of solutions able to supply essential nutrients and buffer pH changes normal to the continued growth of living cells. Such solutions may be complete cell culture media (i.e. any of those commercially available, for instance from Life Technologies Ltd.) optionally supplemented with antibiotics and serum preparations for optimal cell growth conditions. Buffer solutions may also be of the type known as "chemically defined" (e.g. phosphate buffered saline solutions). Cells will also require controlled temperature conditions, in the range 20° to 37°C, and the provision of gases essential to continued cell growth and maintenance of buffer capacity (O₂, and optionally 5% CO₂, depending on the type of buffer being used).

Detection of bioactivity. Detection of bioactivity may be determined by a number of methods known in the art. In a preferred embodiment, detection of bioactivity is determined by cellular imaging of fluorescence. For example, imaging may be conducted of a cell layer on a clear glass substrate. A glass substrate having a surface pitted with a regular array of very shallow (approx 20 µm) depressions may be used for this purpose (Corning). This glass substrate is useful because it ensures a regular and defined spacing between the overlying test array and the cells beneath.

In one embodiment, the detector layer is an undemarcated area of living cells growing on a flat culture surface. The cells on this surface may or may not be grown to confluence, may be transformed and/or engineered cells, or directly derived from animal tissues and grown as primary cell culture. In a second embodiment of the method of the invention, the array of test compounds is laid out onto a non-porous substrate (such as thin coverglass sheet) which is transparent or optically clear. Imaging will be through this surface, and through the cell support membrane lying above. The substrate (Fig. 4, 10) should be inert and solvent tolerant. For example, borosilicate glass sheets of about 200 microns thickness, which may be further surface-treated to give either hydrophobic or hydrophilic properties as desired. This embodiment is shown in the schematic of Fig. 4.

Detector layer: In one embodiment of the invention, the detector layer is a layer of

living cells cultured on a thin porous membrane. A porous membrane useful in the culture and transfer of cells is a transparent non-absorbent membrane with pores of regular and defined diameter which traverse the membrane directly from the upper to the lower side. A porous sheet suitable for cell growth is a track-etched polyester membrane about 10 microns thick with pores between 0.015 and 5 microns diameter at densities of between 600,000,000 to 400,000 pores/cm² respectively (Nucleopore® from Corning Costar).

Delivery of test compounds to detector layer. The porous membrane which supports the detector layer, complete with the buffer medium which overlays it, is applied onto the (dry) test array. Buffer medium wets the lower surface of the porous membrane (Fig. 4, 19) and forms a continuous thin film 23 between the array of test compounds 21 and the porous membrane 19. Test compounds diffuse up through the pores to the detector layer above. In one embodiment of the invention the detector layer is a monolayer of living cells overlayed with physiological buffer solution. The invention includes the possibility that under some conditions it is desirable to have cells grow processes through the membrane to make direct contact with substances on the test array below, with the use of a membrane having an appropriate pore diameter.

Further embodiments and general considerations. Where a test array is generated as a complex mixture of components, such as from the "teabag" method of combinatorial synthesis, or from cDNA library expression systems, a separation step may first necessary. Separation of test components may be conducted in any number of ways known to the art. In one embodiment, components may be separated by the use of one- or two-dimensional separation techniques in non-denaturing gels. The resulting gels may be used directly as test arrays.

Specific separation methods will be tailored to the components involved. Any bioactive compounds from such an array would be identified from identical copies of the original test gel.

Detection of Bioactivity.

Lens and illumination system. Specialized light sources and optics are needed to illuminate and image the fluorescence coming from an area the size of a microtiter plate (96-well plate). Such a system is available from: Imaging Research Inc., St Catherines, Ontario, Canada, and consists of a high-power light source directed through a specialized lens which

acts both as a wide-field epi-illuminator and imaging device.

An illumination system useful in the HDFS device is able to deliver excitation light over an area of at least 8.5 by 13 cm at an intensity sufficient to excite measurable fluorescence from that test field (which in most cases will be living cells loaded with fluorescent reporters). The illumination may come from a scanned beam, or be wide-field for simultaneous illumination of the entire area. The imaging system will collect fluorescent light from the entire test area and bring it to focus onto a sensitive imaging photodetector, such as a cooled CCD camera chip.

Screening. The practice of screening large libraries of samples of unknown composition for the few which may contain a compound of specific biological activity is one of the more common methods of new drug discovery. The samples of unknown composition are in most cases biological material, such as plant extracts or microbial fermentation broths. Screening these for biological activity is normally accomplished by performing binding assays or, more recently, functional assays. A binding assay is an attempt to find compounds of interest by identifying those which adhere with some desired affinity to cells or cell products. This can be done using fluorescent, luminescent, or radioactive detection methods. These assays are based not on a biological response, but passive processes of adherence and displacement. They cannot be construed as functional assays or as real-time assays. Another way to determine biological activity is to measure up-regulation or down-regulation of expression of a known gene. This is done by inserting DNA which codes for something which can be readily measured into a cell's genome such that the expression of interest is coupled to expression of the inserted DNA. While this is a true functional assay, it also is not a real time assay. In addition, it is only capable of finding compounds which affect gene expression. In many cases this is not the response of interest.

The CytoSensor described in U.S. Patent No. 4,915,812 and U.S. Patent No. 5,395,503 is a commercial instrument which has been billed as a screening instrument. It is based on the detection of increased cellular proton flux by means of a semiconducting electrode. The instrument is applicable to high through-put screening, but can only detect cellular events that result in changes in extracellular pH. Again, many responses of interest are not associated with changes in extracellular pH.

The growth over the last few decades in the knowledge of cellular signaling has presented extremely rich opportunities for new ways of screening for biologically active

compounds. Armed with knowledge of the biological process which one wants to affect with a new product, it is possible to monitor the actual process as a way of looking for compounds which affect it. The development of fluorescent probe molecules which upon interaction with intracellular signaling molecules (e.g. ions, enzymes, cyclic nucleotides) change their spectral properties has enabled the real-time monitoring of dynamic biological responses within living cells. Most of these probes can be introduced non-invasively into cells and will, depending on the detection system, allow characterization of cellular events in high temporal resolution (microseconds to seconds) and high spatial resolution (nanometers to micrometers). This probe technology, in combination with the technology of cellular imaging which is described below, has had a major impact on cell biology in that it has enabled monitoring of complex, cross-reacting intracellular events that could not be unravelled by conventional invasive biochemical techniques.

Imaging of cellular functions using luminescent probes. Visualization of intracellular function using luminescent (fluorescent or bioluminescent) probes has become one of the mainstay techniques in modern cell biology. Using traditional optical microscopes with quantitative detectors in place of the human eye, both the concentration and distribution in the cell of a variety of intracellular molecules of interest can be measured. While luminescent probes can be measured in large populations of cells using other techniques, imaging is the only way to learn what is going on in single cells or small populations of cells. The imaging capabilities of the HDFS apparatus will be limited to rather low spatial resolution - fluorescent changes will be imaged from the entire field of detector layer up to 8cm by 12.5 cm. When the detector layer comprises living cells, individual cells need not be resolved in the image, only the fluorescent signals from regions in which cells are present.

The imaging times will vary depending on the responses and parameters being monitored. Signaling responses, for instance changes in the level of free calcium in cellular cytoplasm, may first be seen within seconds or minutes following delivery of test compounds to the detector layer. Such changes can be monitored by changes in the fluorescent properties of specific chemical probes, for instance Fluo-3 or Fura 2 may be used to report on cytoplasmic calcium. The way in which these changes develop within cells (time-response profile) is an important diagnostic feature of the signaling processes giving rise to them. Rapid responses are therefore recorded by sequences of images, where the time between images in a sequence is between 0.1 and 30 seconds (depending on the response being

screened for). Transcription mediated events may require minutes to hours to develop. Monitoring may be continuous or intermittent. For slow responses, two images can be sufficient to gauge the level of response, the first taken before application of test compounds, the second after a period during which the response is estimated to have reached its maximum extent.

Controls relevant to the parameters being measured can be incorporated into the test arrays, both as a check for cell responsiveness and as co-ordinate markers within the arrays. The detector layer is continuous and undemarcated, but because of the close apposition of the test array to the detector layer, the center point of a response in the detector layer corresponds to a conjugate coordinate in the test array. It is helpful to have compounds in the test array which will generate known responses at known coordinates in the detector layer. Responses at the conjugate coordinates in the detector layer act as controls for the system's response, against which responses of the detector layer to unknown compounds may be compared; the points of response to control substances also act as reference points in the detector layer from which the coordinates of other responses can be mapped. For example, when bioactivity is determined as the ability to alter the level of free calcium in cellular cytoplasm, common calcium-mobilizing agonists such as carbamylcholine or adenosine triphosphate are included in the test array at known coordinates.

As another example, when a change in the cellular ratio of inherently fluorescent NAD(P)H/FAD is the biological parameter being assayed, metabolic inhibitors such as KCN or rotenone may be used as a control and marker compounds.

In many instances, diffusion within a thin fluid layer will be involved in many applications of the screening method of the invention, and a concentration gradient will be established from each test point. Those few compounds in a test array which have bioactivity should be detected as spreading rings of response from the focus point of diffusion, within a field of the detector showing no response. The extent of the response areas (measured over time), compared with those from control substances, will provide an indication of potency and solubility of the compound responsible, and also obviate the need to make serial dilutions of test compounds. Toxic or inhibitory substances may also be determined by causing blank sectors in response rings from known agonists. Inhibitory compounds may be determined by their actions on a (pre-)stimulated detector field. Detection of bioactive compounds may incorporate simple image processing to determine the focus, extent and potency/efficacy from

the areas of activity measured in a detector field.

Apparatus

In specific embodiments, the apparatus and method of the invention are as shown in Figs. 1-4. Fig. 1 shows a high energy light source **1**, either a mercury or xenon arc lamp, light from which is collected and collimated by unit **2**, directed through a shutter assembly **3** and passes through an excitation filter-changer **4**. A high-quality light guide **5**, either of fused quartz or a UV-compatible liquid light guide, directs excitation light into the lensing and epi-illumination optics housed in unit **7**. Excitation light emerging from **7** evenly illuminates the horizontal detector layer located in the multi-component assembly labeled **10** and **11**.

Further details of this assembly are shown in Figs. 2a-c, 3, and 4. The assembly comprises two solid layers of which **10** is fixed relative to the stage unit **8** which supports it, while layer **11** is moved vertically downward on guide pins (**17** in Figs. 2a,b,c) to bring test compounds into contact with the detector layer. Vertical movement of **11** is controlled by arm **12** driven by unit **13**. Four sprung contacts **14** attached to **12** press upon the frame of layer **11** to drive it downwards as arm **12** descends. A separate drive unit **9** controls position of the stage **8** in the horizontal plane, and also is used to adjust focus by movement along the vertical axis.

Fluorescent light emitted by the detector layer is collected by lensing unit **7**, passes through an emission filter-changer **15** and is brought to focus on the photosensitive surface of an imaging detector housed in unit **16**.

Specified devices (**3, 4, 9, 13, 15, 16**) are controlled by a central processing unit **6** which issues commands to, and collects data and status information from the devices attached to it. Collected data (images) can also be analyzed by unit **6**, or passed to a subsidiary analysis station (not shown). Unit **6** comprises: central processing unit (Intel Pentium chip, or better), RAM, multi-channel serial input/output cards with onboard A/D and D/A converters, one of which cards controls the camera **16** and captures images from it, also a video controller card, VDU, and hard disk memory units.

Figs. 2a,b,c are schematic diagrams of the test stage, which includes a supporting stage **8** with large rectangular central aperture, the shape and size of which is the same as the area labeled **19**. The position of stage **8** is adjusted in the horizontal and vertical axes by the

3-axis positioner **9**. These diagrams are drawn for the specific embodiment in which the detector layer is a layer of living cells growing on the upper surface of the solid transparent component **10**, which also serves to contain the liquid layer **18** which overlays the cells in the detector layer and provides them with necessary nutrients and conditions to keep them alive. The printed array of test compounds **21** is borne on a sheet of track-etched membrane **19** held by a rectangular rigid frame **11**. At the beginning of the screening assay, the array **21** is not in contact with the fluid layer **18**. The array **21** is held away from the liquid layer by pins **17** which pass through holes **24** in the corners of the frame **11** and which, by friction or "click-stops", prevent it from falling (Fig. 2a). At the appropriate moment, arm **12** is moved down by the drive unit **13** and the four sprung contacts it bears **14** exert pressure on the frame **11** moving it down the guide pins **14** and into the liquid **18** below to a position where it is in very close proximity to the underlying layer of detector cells **20** grown on top of the solid substrate **10** (Fig. 2b). Throughout this procedure, the entire area of the detector layer corresponding to the size and shape of area **19** is illuminated and imaged from below by the additional apparatus shown in Fig. 1.

The apparatus can also be used in a second embodiment of the screening method of the invention, where the test array is laid out on the upper surface of component **10**, and components **11** and **19** are a frame and thin transparent track-etched membrane, respectively. In this specific embodiment, the frame **11** is sufficiently deep to contain culture liquid as required to sustain the detector layer of living cells growing on the upper surface of the membrane **19**.

Figs. 3 and 4 are schematics to show the relative positions of the different layers in the test-array/detector layers used in the specific embodiments of the invention. Fig. 3 shows the arrangement in which an array of discrete spots of test compounds **21** on a porous membrane **19** is in contact with a liquid layer **18** overlaying the detector layer **20** which is supported by an optically transparent solid substrate **10**. The compounds fill the parallel capillary spaces **22** in the track-etched membrane **19**. Bioactivity is detected by measuring changes in fluorescence in the detector layer **20** resulting from responses to the diffusion of test compounds through the porous membrane to the detector layer.

Fig. 4 is a schematic drawing of a second embodiment of the screening method in which a detector layer **20** supported on an optically clear porous membrane **19**, and overlayed

by a liquid layer 23, is placed onto an optically clear solid substrate 10 bearing an array of test compounds 21. The thin space 18 between components 19 and 10 is filled with solution from 23 which has passed through the porous membrane 19. Bioactivity is again detected by measuring changes in fluorescence of the detector layer resulting from responses to the diffusion of test compounds through the porous membrane to the detector layer.

Fig. 5 is a schematic illustrating the way in which an array of 1536 compounds can be created on a membrane surface, such as would be useful in the first embodiment described above, by simple transfer printing. Compounds are stored in 16 separate 96-well microtiter plates and defined amounts are transferred simultaneously by a 96-pin printing head to the surface 19. The contents of each successive 96-well plate are printed at a slightly offset position, generating an array as shown in Fig. 5b after 4 such printing operations, and a full array of 1536 compounds (Fig. 5c) after 16 printing operations. The holes 24 in frame 11 are used to position and guide the completed array on the pins 17 indicated in Figs. 2b and 2c. The process illustrated in Fig. 5 can also be used to transfer an array of test compounds to a solid surface such as would be useful for component 10 in the second embodiment of the method described above.

EXAMPLE

Example 1. Screening of 1536 Test Compounds for Bioactivity.

The following description of the use of one embodiment of the apparatus of the invention in the screening method disclosed. An array of test compounds are supplied in 96-well microtiter plates, as is common practice for compounds produced by methods commonly known as combinatorial chemistry, or for compounds extracted from natural sources. In this example, the compounds are provided in soluble form, and the concentrations and solvents used have previously been tested for compatibility with the apparatus. In this example, 1536 compounds are tested simultaneously against a known cellular target, specifically a G-protein coupled receptor (GPCR) of the Gq type expressed in a transformed cell line. Gq GPCRs give clearly identifiable changes in intracellular calcium when activated.

First, physiologically viable living cells are cultured to a near confluent monolayer in a transparent culture dish (10, Fig. 2a-c) in appropriate culture medium and conditions. Immediately prior to being used in the experiment, the cells are loaded with the fluorescent

indicator of free cytoplasmic calcium concentration, Fluo-3 (from Molecular Probes, Oregon). This is accomplished by incubating the cells with a 2 to 5 μ M solution of Fluo-3 acetoxymethyl ester (AM) for a period of 10 to 15 minutes, followed by a series of solution exchanges to wash away excess Fluo-3 AM.

The method of transfer of compounds to the track-etched membrane Fig. 2a-c **19** is illustrated in Fig. 5. In this example, 1536 compounds are printed as an array **21** on a single track-etched membrane **19**, from sixteen individual 96-well microtiter plates in the following manner: A 96-pin printing head is used to transfer defined volumes of compounds (in the range 0.05 to 0.5 μ l of each compound), one compound per pin, from each 96-well plate in turn (with wash steps between source plates to avoid cross-contamination). Each 96-point print to the membrane occurs in an offset grid, such that 16 print operations are made sequentially on the same membrane and the printed spots of compounds remain discrete and separated from each other (three of these spots are indicated in Fig. 5a, **21**). Fig. 5a shows the result of a single 96-point print operation, Fig. 5b after four such operations, and Fig. 5c the finished array after 16 print operations. In this way, just sixteen print operations (and sixteen intermediate wash steps for a single print head) are sufficient to transfer 1536 compounds to a single test array. The procedure can be readily automated, and multiple copies of each printed sheet made for multiple tests.

Completed arrays are fixed to the pins **17** (Figs. 2b-c) projecting from the culture dish **10** such that they are supported some small distance above the thin fluid layer **18** covering the living cells which form the detector layer. Once the test array is fixed in place over the Fluo-3-loaded cells, the entire assembly is placed onto the test stage as shown in Fig. 2a.

The following events are synchronized by sequential instructions from the computer processing unit **6**. First, the test stage is centered over the lensing unit **7** (Fig. 1) and the detector layer it supports is brought into focus by the motor unit **9**. Fluo-3 is excited by light of 490 nm, and its fluorescent emissions are collected in the range 505-540 nm. The intensity of emission is increased when the dye binds free calcium. Thus the computer brings a 490 nm band-pass excitation filter into line of the light path coming from units **1** and **2** using the filter changer unit **4**. At the same time, a band-pass emission filter for the range 505-540 nm is positioned in the imaging path by unit **15**. The shutter **3** is opened for a pre-determined exposure period (typically 50 to 500 milliseconds), and during this time the whole area of the

detector layer is illuminated with 490 nm light. Fluorescent emission from the Fluo-3 in the cells is collected by the lens 7 and focused into the camera. The camera captures the image and sends it to the processing unit 6 where it is stored and displayed. At regular intervals thereafter, images are captured in sequence by repeatedly opening the shutter 3. Intervals between successive images are typically in the range 0.5 to 30 seconds, depending on the speed of the response expected. Intervals of 0.5 to 2 seconds are usual and sufficient to sample the dynamics of most changes in cellular calcium. At a predetermined time during this continuing sequence of images, the test array is pushed down the guide pins 17 by the actuating arm 12 and its sprung contacts 14, driven by unit 13. In close apposition to the cells in the detector layer, the test array begins to release the compounds it carries. The compounds dissolve into the liquid layer, and these fall vertically downwards onto the cells below. Because there is only a thin liquid layer between the membrane of the test array and the cells below, there is insignificant intermixing of adjacent test compounds. If a test compound activates cells below it bearing Gq GPCRs, these cells will respond with an immediate increase in free cytoplasmic calcium, and the fluorescence signal from the Fluo-3 dye they contain will increase. The sequence of images collected during the period of the response (which is typically of 1 to 10 minutes duration) will reveal which cells have so responded, and their position in the area of the detector layer will be correlated with the identity of the compound in the array above. An analysis of the entire area of each image in the sequence, performed on-line by the processing unit 6, yields the following information: the identity of the compound eliciting the response, the profile of the response with time, the intensity of the response, and also the potency of the compound with reference to a chosen standard. The latter information is contained in the radius of the area of cells responding within a particular time, and can be compared directly to a known standard which is included in the array at known points. The use of standard compounds at known points in the array also provides a control for the experiment, and helps to identify coordinates in the detector layer from which other responses can be mapped.

At the end of the screening assay, the sequence of images is stopped, the actuating arm 12 raised, and the test assembly removed. The next assembly is then moved in and the sequence begun afresh. Assembling the test units and exchanging them on the test stage can be automated by appropriate robotic control (not shown in the diagrams).

One of the advantages of the method of the invention is that the method does not require that either the components of the detector layer (e.g. living cells), or the different test compounds, be isolated from one another within discrete chambers or compartments, as is common to all high throughput screening procedures currently in use or development. The method also removes the need to dispense microvolumes of test compounds during the period of the assay itself. Delivery of test compounds to detector layers is either by direct contact or by simple diffusion across thin liquid films. Delivery and detection becomes a (massively) parallel process.

CLAIMS

What is claimed is:

1. A method for screening test compounds for bioactivity, comprising:
 - (a) contacting an array of test compounds with a detector layer; and
 - (b) detecting a detector layer response, wherein a response is indicative of bioactivity.
2. The method of claim 1, wherein the detector layer is comprised of physiologically viable cells.
3. The method of claim 2, wherein the detector layer is supported by an optically clear substrate.
4. The method of claim 3, wherein the reactive sensing surface is held stationary in the field of view of the optical detector and the sample surface is moved into contact with it during the course of the measurement.
4. The method of claim 1, wherein the detection of step (b) is a change in a fluorescence or luminescence property of the cell.
5. The method of claim 4, wherein detection is determined with an illumination system capable of exciting the fluorescence of the reactive surface with any of a number of previously selected wavelengths with defined order and of defined time duration.
6. The method of claim 2, wherein the physiologically viable cells form a monolayer.
7. The method of claim 1, wherein the test compounds are generated on a solid support by combinatorial chemistry.
8. The method of claim 1, wherein the test compound array is generated by one- or two-dimensional gel electrophoresis.

9. A method for high throughput screening of test compounds for bioactivity, comprising:

(a) contacting a solid support comprising an array of multiple test compounds with a cell layer, wherein each test compound comes into contact with a localized liquid which is in contact with a detector layer; and

(b) detecting a response of the detector layer to the test compound, wherein a response is indicative of a bioactive compound.

10. A method for simultaneously exposing an array of test compounds with a reactive sensing surface, comprising the steps of:

(a) contacting an array of test compounds on a solid matrix with a porous membrane which is in contact with a liquid layer overlaying a reactive sensing surface layer; and

(b) allowing the test compounds to diffuse through the porous membrane to the liquid layer overlaying the reactive sensing surface.

11. An apparatus for screening an array of test compounds for bioactivity, comprising:

(a) a solid support comprising an array of test compounds;

(b) a porous membrane; and

(c) a detector layer layer, wherein a liquid layer is between the porous membrane and detector layer layer, and wherein the test compounds contact the detector layer layer by diffusion through the porous membrane.

METHOD AND APPARATUS FOR HIGH DENSITY FORMAT SCREENING FOR BIOACTIVE MOLECULES

Abstract

A method and apparatus for screening an array of test compounds for bioactivity by contacting an array of test compounds with a detector layer capable of detecting bioactivity, and detecting a detector layer response. The detector layer is comprised of physiologically viable cells. The method and apparatus allow a large number of test compounds to be simultaneously assayed in parallel.

Fig. 1

Schematic view of equipment; not to scale

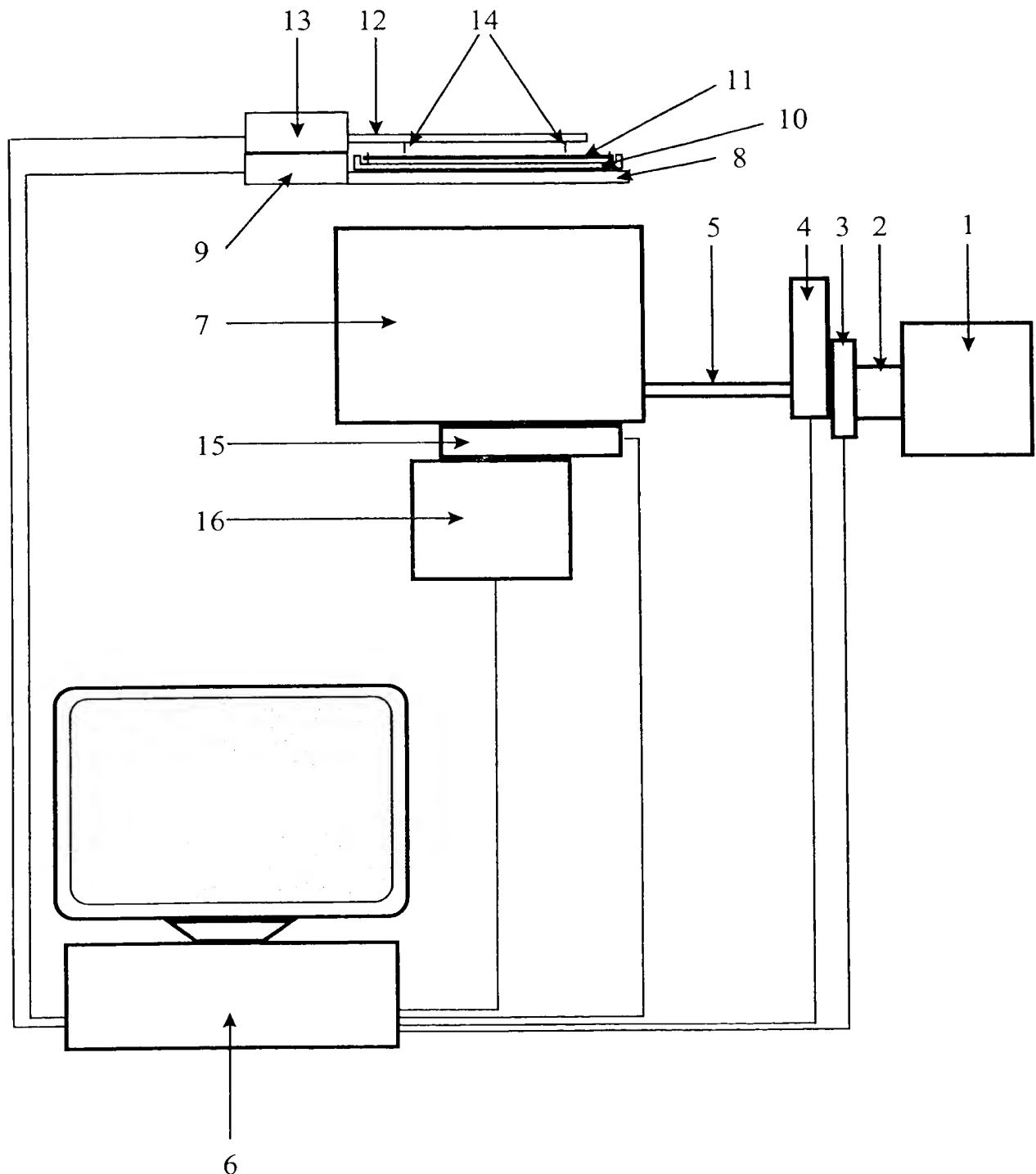
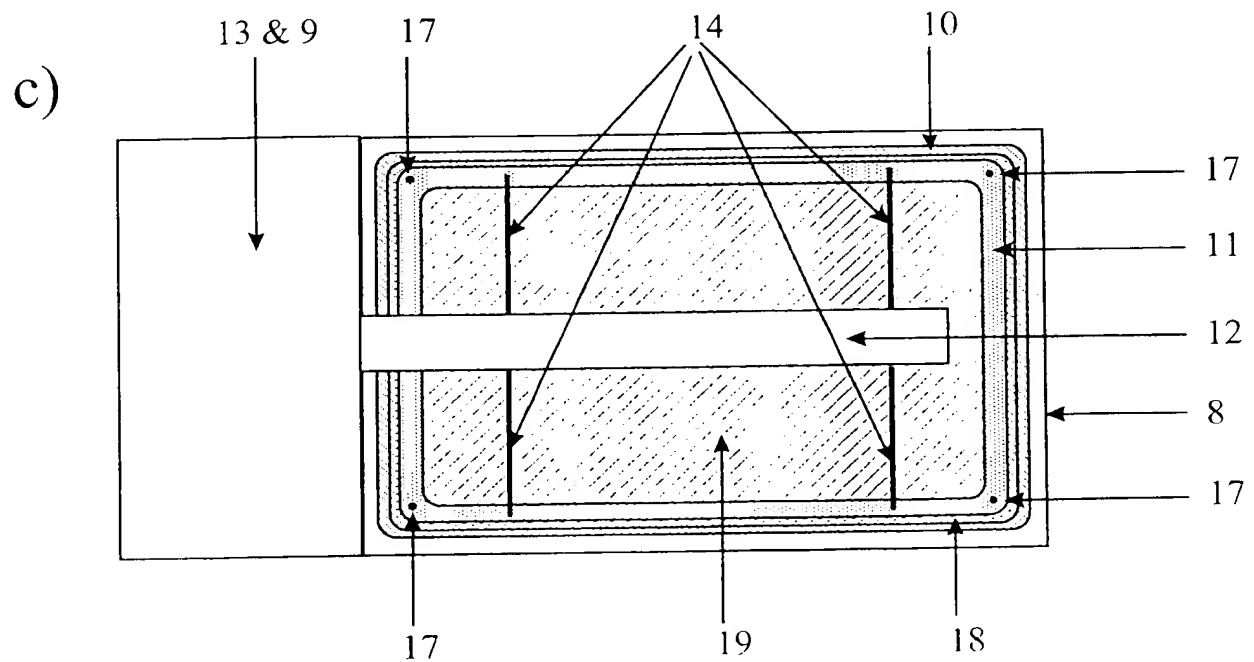
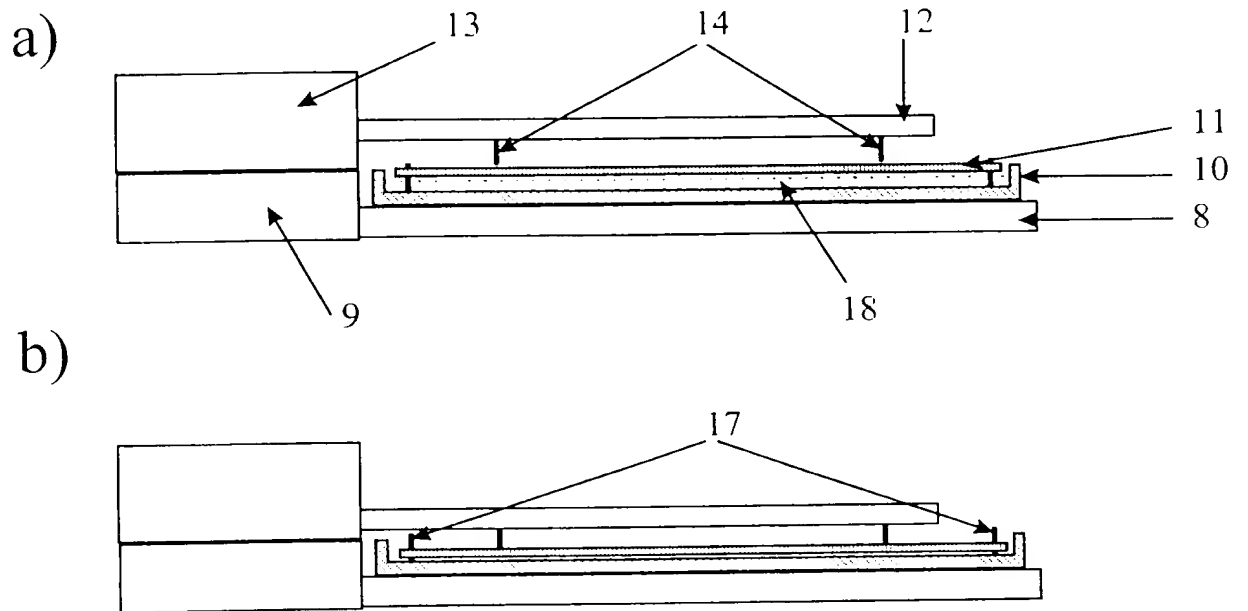


Fig. 2

Side views of test stage; not to scale



Top view of test stage; not to scale

3-D sectional representations of portions of
the test-array/detector layers: not to scale

Fig. 3

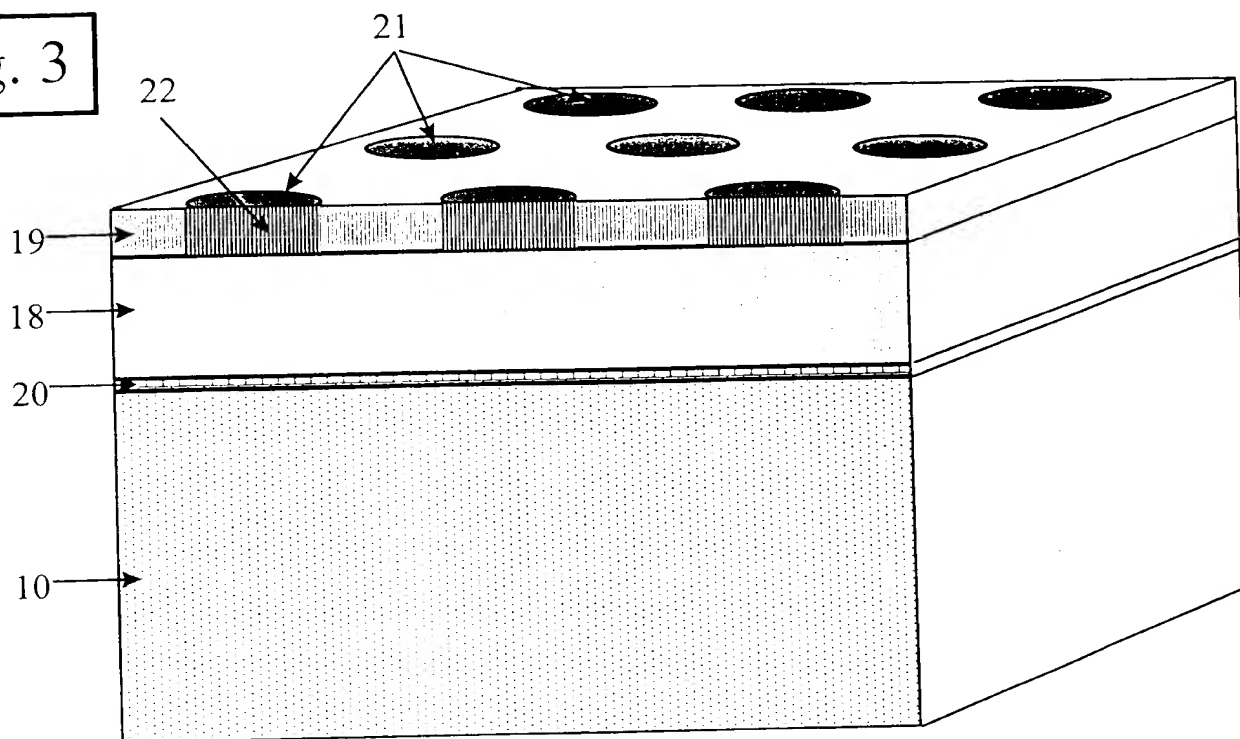
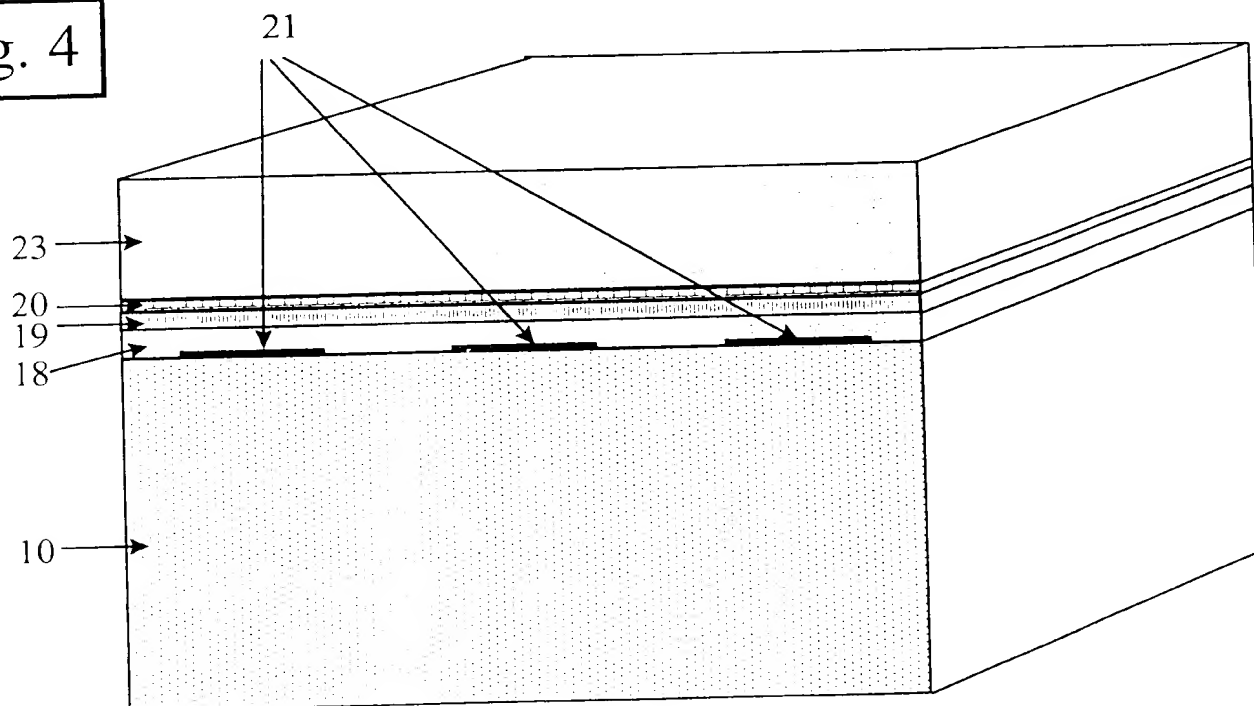


Fig. 4



Changes in intracellular cAMP visualised using a cAMP-dependent protein kinase-green fluorescent protein hybrid.

Kasper Almholt, Bernard R. Terry*, Ole Skyggebjerg, Søren Tullin, Kurt Scudder, and Ole Thastrup

BiolImage, Novo Nordisk A/S, 28 Mørkhøj Bygade, DK-2860 Søborg.

Keywords: PKA, cAK, C-subunit, cAMP, measurement, live-cell, GFP, redistribution

*corresponding author:

Tel.: +45 4442 6341

Fax: +45 4442 1411

e-mail: bobt@novo.dk

ABSTRACT

A novel method to monitor changes in intracellular cAMP concentration ([cAMP]_i) within intact living cells has been developed based on a fusion of the catalytic subunit of cAMP-dependent protein kinase to green fluorescent protein (GFP). In stably transfected unstimulated fibroblasts, fusion protein fluorescence was highly concentrated in aggregates throughout the cytoplasm and absent in the nucleus. Stimulation with the adenylate cyclase activator forskolin caused the release of tagged catalytic subunits from the cytoplasmic aggregates within minutes, resulting in an increasingly homogeneous distribution of GFP fluorescence throughout the cytoplasm. The observed redistribution was completely reversible: removal of forskolin led to the return of fluorescence to the cytoplasmic aggregates. Spot-photobleach measurements showed that the rate of exchange of GFP-labelled catalytic subunits at these aggregates increased in proportion to [cAMP]_i. The localisation of the fusion protein was also sensitive to receptor stimulation. In fibroblasts stably expressing the G_s-protein coupled glucagon receptor, generation of an increased [cAMP]_i through glucagon stimulation resulted in a redistribution of tagged catalytic subunit similar to that observed after forskolin addition. Conversely, in fibroblasts overexpressing the G_i-protein coupled α 2a adrenoreceptor, addition of norepinephrine after forskolin stimulation led to a reversal of the fusion protein redistribution.

INTRODUCTION

The cAMP-dependent protein kinase (cAK)¹ is a ubiquitous serine/threonine protein kinase. cAK is recognised as the only mediator of intracellular cAMP signals in eukaryotes², with the exception of certain ion channels³. The cAK holoenzyme is an R₂C₂ tetramer consisting of a regulatory (R) dimer and two catalytic (C) subunits². Presently, four isoforms of the regulatory subunit (RI α , RI β , RII α and RII β) and three isoforms of the catalytic subunit (C α , C β and C γ) have been described². Splice variants of C α and C β ⁴ and possible R heterodimers, as reported for RI α and RI β ⁵, add to the complexity of the cAK holoenzyme. Although the C γ isoform is unique with respect to substrate specificity, inhibition and tissue distribution⁶, few reports suggest different roles for C α and C β isoforms of the catalytic subunit⁷. In contrast, the RI and RII subunits are reported to be distinct. The cAKI (RI₂C₂) holoenzyme is thought to be mainly soluble and cytoplasmic² although RI is reported to be associated with

sarcoplasmic membranes⁸ and also with a detergent-resistant structure in mammalian sperm⁹. cAKII (RII₂C₂) on the other hand is thought to be particulate and RII has been reported to bind to a number of intracellular components, most notably Golgi membranes^{10,11} and centrosomes^{10,11} but also mitochondria¹², nuclei^{13,14} and cytoskeletal components^{11,12}. RII subunits interact with a family of proteins called A-kinase anchoring proteins (AKAP)¹⁵ and this may also be true of RI subunits¹⁶. The AKAP-RII subunit interaction is presumed to be responsible for localising the cAKII tetramer at these intracellular sites. The NH₂-terminus of the C subunit is myristoylated¹⁷, a post-translational modification usually associated with membrane insertion. However, the C subunit does not appear to be membrane attached and while myristoylation may increase the thermostability of the protein, the possible role of myristoylation in its targeting or substrate specificity is still not clear¹⁸.

The C subunit in the assembled tetramer is believed, although not unanimously¹⁹, to be catalytically inactive. Activation of cAK is physiologically mediated through G_s-protein coupled plasma membrane receptors. G_s-protein activation leads to activation of adenylate cyclases, which generate cAMP. Binding of two molecules of cAMP to each R subunit causes the release and activation of the C subunits. Dissociated C subunits phosphorylate cytoplasmic substrates^{20,21} and have been shown to relocate to the nucleus²². The nuclear redistribution mechanism of C subunits may be by simple diffusion through nuclear pores²¹. To date a large number of cytoplasmic and a few nuclear cAK substrates have been reported. An incomplete list of 25 *in vitro* substrates²³ includes several enzymes involved in basic metabolism such as phosphorylase kinase, glycogen synthase and fructose biphosphatase. Nuclear C subunit regulates transcription of genes under control of the cAMP response element (CRE) by phosphorylating the continuously bound CRE binding protein, (CREB)^{24,25}.

Several factors decrease the level of cAK activity. Stimulation of plasma membrane bound G_i-protein coupled receptors inhibits adenylate cyclases and cAMP is continuously being broken down by a variety of phosphodiesterases. Despite the importance of the cAMP/cAK signalling pathway, there is no easy method to monitor intracellular cAMP concentrations ([cAMP]) in intact living cells. The current method of choice involves fluorescence resonance energy transfer (FRET) between microinjected fluorescently labelled R and C subunits²⁶. In the work described herein, the C α subunit was tagged with a highly fluorescent variant of green fluorescent protein (GFP) containing F64L and S65T amino acid substitutions (GFP^{L1}) (International

Publication No. WO97/11094). This approach provides a transfectable probe for monitoring the intracellular trafficking of C subunits in response to changes in [cAMP], and represents the first easy method to evaluate changes in [cAMP], in intact living cells in response to extracellular signals.

Results

GFP^{LT} tagged C had the expected molecular weight.

Lysates of glucagon receptor-transfected baby hamster kidney cells (BHK/GR) stably expressing the C-GFP^{LT} fusion protein were characterised by Western blot analysis using polyclonal antibodies directed against the NH₂-terminus of C α (Fig. 1). In a separate experiment, lysates of BHK cells, transiently expressing either of the two fusion proteins, were characterised by Western blot analysis using polyclonal antibodies that recognise GFP (data not shown). Taken together, these experiments show that C-GFP^{LT} fusion protein is recognised as a unique protein of the expected size by the anti-C α antibody in stably transfected cells and that both fusion proteins have the same molecular weight.

The fusion protein localised to cytoplasmic aggregates.

The localisation of the two fusion proteins, when transiently expressed in Chinese hamster ovary (CHO) cells, was very different. While GFP^{LT}-C was evenly distributed throughout the cytoplasm (Fig. 2A), C-GFP^{LT} was found in highly fluorescent aggregates in the cytoplasm (Fig. 2B). These distinct patterns for the two fusions was also seen in transiently transfected human embryonic kidney (HEK293) and BHK/GR cells (data not shown). For unknown reasons it was not possible to make stable transfectants expressing the GFP^{LT}-C fusion, whereas this procedure was straightforward with the C-GFP^{LT} fusion. The distribution of GFP^{LT}-C in transiently transfected CHO cells did not change when [cAMP] was raised by the addition of 50 μ M forskolin (n=6, data not shown). The following results are therefore based only on work with the C-GFP^{LT} fusion.

Increased [cAMP]_i caused the release of fusion protein from cytoplasmic aggregates.

Within 2-3 minutes of treatment of CHO/C-GFP^{LT} cells with forskolin, C-GFP^{LT} fluorescence dispersed from the bright aggregates and filled the cytoplasm (Fig. 3A, 1 μ M forskolin), remaining in this distribution for as long as forskolin was present (cells were followed up to two hours). The probe did not enter the nuclear compartment to any clearly observable extent. Higher doses of forskolin increased the rate and extent of probe redistribution. The responses depicted in Figure 3B-G have all been quantified from image data, as described in the experimental protocol. Table 1 gives a comparison of the average temporal profiles of fusion protein redistribution in response to the three forskolin concentrations shown in Figure 3B. Addition of 1 mM dibutyryl cAMP (dbcAMP) (n=6), a membrane permeable cAMP analogue, which is not degraded by phosphodiesterases, caused a similar but slower response (Fig. 3C). Addition of 100 μ M 3-isobutyl-1-methylxanthine (IBMX) (n=4), a cell permeable phosphodiesterase inhibitor, caused a similar, slow response (Fig. 3D), even in the absence of adenylate cyclase stimulation. Addition of buffer (n=2) had no effect (data not shown). As a control for the behaviour of the fusion protein, GFP^{LT} alone was expressed in CHO cells and these also given 50 μ M forskolin (n=5); the uniform diffuse distribution characteristic of GFP in these cells was unaffected by such treatment (data not shown).

To test the reversibility of the fusion protein redistribution, CHO/C-GFP^{LT} cells were treated with 10 μ M forskolin (n=2) and washed repeatedly (5-8 times) with 37°C buffer. Although the plant terpenoid forskolin is lipophilic, it is possible to remove its effect by washing with aqueous buffer²². In these experiments, fusion protein began to return to its prestimulatory localisation within 2-3 min (Fig. 3E). In fact the fusion protein returned to a pattern of fluorescent cytoplasmic aggregates virtually indistinguishable from that observed before forskolin stimulation. To test whether the return of fusion protein to the cytoplasmic aggregates reflected a decreased [cAMP]_i, cells were treated with a combination of 10 μ M forskolin and 100 μ M IBMX (n=2); when washed repeatedly (5-8 times) with 37°C buffer containing 100 μ M IBMX the fusion protein did not return to its prestimulatory localisation after removal of forskolin (Fig. 3E).

To test the probe's response to receptor activation of adenylate cyclase, stably transfected BHK/GR,C-GFP^{LT} cells were exposed to glucagon stimulation. In these cells, addition of 100 nM glucagon (n=2) caused the release of C-GFP^{LT} from the cytoplasmic aggregates and a resulting permanent redistribution of the fusion protein to

a more even cytoplasmic distribution within 2-3 min (Fig. 3F). Similar but less pronounced effects were seen at lower glucagon concentrations ($n=2$, data not shown). Addition of buffer ($n=2$) had no effect over time (data not shown). CHO/C-GFP^{LT} cells, transiently transfected with the $\alpha 2a$ adrenoreceptor (AR $\alpha 2a$), were treated with 10 μ M forskolin then, in the continued presence of forskolin, exposed to 10 μ M norepinephrine to stimulate the exogenous adrenoreceptors. This treatment led to reaggregation of C-GFP^{LT} within the fluorescent structures, consistent with a receptor-induced decrease in [cAMP]_i (Fig. 3G).

Rate of recovery from photobleach of C-GFP^{LT} aggregates is dependent on forskolin concentration.

Photobleach measurements were made to confirm that changes seen in the distribution of C-GFP^{LT} fluorescence were a result of changes in the rate of turnover of C-GFP^{LT} upon the aggregates. The fluorescence of an entire C-GFP^{LT} aggregate within a cell could be effectively bleached within 2 to 5 seconds by a stationary laser beam at full intensity. After bleaching, aggregates recovered their fluorescence, indicating a dynamic exchange of C-GFP^{LT} at these loci (Fig. 4A). The rate of recovery from spot photobleach was highly reproducible at each particular concentration of forskolin even in different cells (Fig. 4B). Both the extent and rate of recovery increased with the forskolin treatment given. Most recovery curves required at least two exponentials to fit them adequately. Given the limits of the experimental procedure, the curves are used here only to estimate half-times of recovery. To an approximation, half times for recovery can be estimated directly from the slope of reciprocal plots of the fluorescence displacement for the first few time points²⁷. Values for half times estimated within the first 3.0 seconds of recovery (Fig. 4C) are plotted as a dose response curve in Figure 5, giving an estimated $\frac{1}{2}$ -maximal concentration for forskolin of about 3 μ M

Fusion protein redistribution correlated with [cAMP]_i

As described above, the time it took for a response to come to completion was inversely related to the forskolin dose (Table 1). In addition the extent of a response was also dose dependent. In an automated imaging system we stimulated CHO/C-GFP^{LT} cells with 5 increasing doses of forskolin ($n=8$). Images were analysed with the same algorithm used

to construct Figure 3B-G. From the results shown in Figure 5, a half maximal stimulation was observed at 1.7 μ M forskolin by this method. In parallel, CHO/C-GFP^{LT} cells were stimulated with 8 increasing concentrations of forskolin (n=N) and the relative amount of cAMP produced was measured in a scintillation proximity assay (SPA). The 1/2-maximal concentration for forskolin in the SPA assay was determined to be 9.3 μ M (Fig. 5).

Co-localisation of C-GFP^{LT} with labelled ceramide distributions

Figure 6A is an overlay of green and red fluorescence emissions from CHO/C-GFP^{LT} cells stained with BODIPY[®] FL C₅-ceramide (ceramide-FL). The green channel contains the ceramide-FL and GFP^{LT} fluorescence; the red channel shows only the ceramide-FL excimer emission. The ceramide-FL probe preferentially accumulates in Golgi membranes²⁸. This is most obvious in images formed from the red excimer emissions of the FL-ceramide. The GFP^{LT}-bright structures do not stain with the ceramide probe indicating that they are clearly distinct from Golgi membranes.

Structure of the GFP^{LT} -bright aggregates

Figure 6B shows an iso-surface rendering of 25 deconvolved and reconstructed through-focus wide-field images of a single large C-GFP^{LT} aggregate. Each aggregate appears to have the structure of a convoluted tubule or glomerulus, and this is more obvious in the stereo pair (Fig. 6C) derived from the same data set from which the iso-surface rendering was made. It is not completely clear whether each structure is formed from a single fully connected tubule or a small number of discrete tubules in close apposition. The structure is however clearly compact and more complex and structured than a simple amorphous aggregation of C-GFP^{LT} molecules. Figure 6B-C is typical of the larger aggregates which are of the order of 2 to 4 μ m across. The more numerous smaller aggregates (less than 1 μ m across) appear to share the same underlying structural component(s) as their larger counterparts.

Discussion

The aim of the present study was to develop a transfectable probe for monitoring changes in $[cAMP]_i$. Since cAK is by far the major intracellular effector for $cAMP^2$, a measure of its activation should closely reflect physiologically relevant changes in $[cAMP]_i$.

NH_2 - and $COOH$ -terminal fusions of C subunit were made to a highly fluorescent variant of GFP. Only the C-GFP^{LT} fusion responded to changes in $[cAMP]_i$. The three-dimensional structure of the C subunit^{29,30} reveals that both the NH_2 - and $COOH$ -termini, while far apart, are both located opposite the catalytic cleft and close to the surface of the protein. Comparison with the closely related cGMP-dependent protein kinase, whose R and C subdomains are contained within the same polypeptide chain in R-C order³¹, suggests that the R subunit of cAK may be expected to interact with the NH_2 -terminal region of the C subunit. Furthermore, the surface of the C subunit in the NH_2 -terminal region is hydrophobic²⁹, supportive of a protein-protein interaction in this area. An NH_2 -terminal GFP^{LT} tag would also prevent post-translational myristoylation (of the NH_2 -terminus) of the C subunit as reported specifically for mouse $C\alpha^{18}$, while the C-GFP^{LT} fusion may well be myristoylated. These factors may explain the very different behaviours of the NH_2 - and $COOH$ -terminal fusions of C subunit to GFP^{LT}.

There are reasons to believe, that the C-GFP^{LT} fusion protein behaves like the endogenous kinase both with regard to localisation and activation kinetics. Li *et al.* (1996)¹¹ have, for instance, reported that RII subunits occur as "intensely fluorescent spots" within perinuclear cytoplasm. Skålhegg *et al.* (1997)³² also reported a granular distribution of RII in both human B and T lymphocytes. Also, the time frame of fusion protein redistribution in response to forskolin addition reported here, corresponds well to the observation of dissociation of microinjected $RI\alpha_2C\alpha_2$ holoenzyme in response to forskolin within 1-2 minutes²⁶ and the dissociation of endogenous RII_2C_2 in response to forskolin observed by immunofluorescence after less than 5 min²².

In contrast with previous work with microinjected $RII\alpha_2C\alpha_2$ holoenzyme and $C\alpha$ subunit²¹, we did not observe any translocation of C-GFP^{LT} to the nucleus. A possible explanation could be the increased size of the fusion protein relative to endogenous C subunit. Nuclear pores are thought to allow passage by diffusion of globular proteins of less than 45-60 kDa³³. The putative size limit of 45-60 kDa may adequately explain the exclusion of the fusion protein (68 kDa), yet passage of endogenous C subunit (41 kDa).

Consistent with this, a microinjected 65 kDa fusion protein of glutathione S-transferase and mouse C α subunit (GST-C) was excluded from the nucleus²¹.

That the C-GFP^{LT} fusion can be released by dbcAMP or treatments which increase [cAMP]_i suggests that it must recognise and attach to endogenous R subunits (or some subset of the same) and therefore that these R subunits are naturally collected at or on the structures seen in Figures 3A and 6. Reversal of elevated [cAMP]_i, e.g. by removal of forskolin or stimulation of G_i-coupled receptors, results in rapid return of fluorescence to the original prestimulatory locations within cytoplasm. These anchoring structures therefore appear to be persistent features within the cytoplasm of CHO/C-GFP^{LT} cells. Similar structures and C-GFP^{LT} behaviour were also found in transfected BHK cells.

The distribution of fluorescence between aggregates and cytoplasm should reflect the position of a dynamic equilibrium within each cell, determined principally by [cAMP]_i. This is confirmed by results from spot-photobleach measurements. The rate of fluorescence recovery of aggregates following photobleach measures the net rate of turnover of C subunits at these sites. The rate of recovery is the sum of on and off rates for the association of catalytic with regulatory subunits at these loci, both of which will be governed principally by the concentration of cAMP within the cell (the off rate being governed directly by [cAMP]_i; the on rate being dependent on the concentration of free C-GFP^{LT} in the cytoplasm). Most aggregates completely disappear after full stimulation with forskolin. However, often one aggregate remains, and this is always the biggest and brightest from the unstimulated cell. Nevertheless, as photobleaching can demonstrate, there is active turnover of C-GFP^{LT} even at these large fluorescent aggregates which remain in fully stimulated cells. As a further observation, there appears to be considerable mobility of catalytic subunits within the structure of an aggregate, since a stationary laser beam (approx. 0.5-1.0 μ m diameter) is able to bleach fluorescence from an entire aggregate of 2-3 μ m diameter in 2 to 5 seconds.

The lack of colocalisation of C-GFP^{LT} and ceramide fluorescence, the position of aggregates within the cell and their unusual form, suggest that these structures are definitely not associated with Golgi, but may well be constructed of membrane tubules with C-GFP^{LT} on the outer surface. Although we have been unable as yet to ascertain the identity of these structures, we have ruled out Golgi membranes. They may however be membranous since fusion protein is apparently freely mobile on them, possible tubular judging by the 3-D reconstructed image, and clearly the catalytic subunits are able to

bind to and release from R subunits with ease, suggesting that the latter are anchored to the surface of these structures. They are also persistent within the cytoplasm, and found in all cells transfected thus far with the C-GFP^{LT} construct (CHO, HEK293 and BHK).

Figure 5 gives a comparison of an SPA assay conducted in parallel with two different forskolin dose response experiments using the cAK fusion protein. These experiments showed a direct correlation of three parameters: level of [cAMP]_i, turnover rate of C-GFP^{LT} at cytoplasmic aggregates, and overall degree of fusion protein redistribution. Data from these three greatly varying methods agree on an 1/2-maximal concentration for forskolin of between 1.7 to 9.3 μ M in this system. As these results show, the cAK fusion protein represents a novel and reliable probe by which dynamic changes in [cAMP]_i can be measured in intact living cells as they respond to extracellular signals.

Experimental protocol

Hybrid cDNA construction

Hybrid cDNAs encoding NH₂- and COOH-terminal fusions of murine C α subunit³⁴ to GFP^{LT} were inserted into the multiple cloning site of the pZeoSV (Invitrogen Corp., San Diego, CA, USA) mammalian expression vector, generating the fusion constructs C-GFP^{LT} and GFP^{LT}-C. Briefly, cDNAs encoding C and GFP^{LT} were amplified by PCR using the following primers:

5'-C, TTGGACACAAGCTTTGGACACCCTCAGGATATGGGCAACGCCGCCGCCGCC AAG;
 3'-C, GTCATCTTCTCGAGTCTTTCAGGCGCGCCCAAACCTCAGTAAACTCCTTGCCA CAC
 ;
 5'-GFP^{LT}, TTGGACACAAGCTTTGGACACGGCGCGCCCATGAGTAAAGGAGAAGAAGCTTT TC
 and
 3'-GFP^{LT}, GTCATCTTCTCGAGTCTTACTCCTGAGGTTTGTATAGTTCATCCATGCCATGT .

HindIII/AscI restriction endonuclease digested C subunit PCR amplification product and AscI/XhoI digested GFP^{LT} PCR product were ligated with the HindIII/XhoI digested vector for the generation of the C-GFP^{LT} fusion construct. Correspondingly the GFP^{LT}-C construct was generated by ligating HindIII/Bsu36I digested GFP^{LT} PCR product and Bsu36I/XhoI digested C subunit PCR product with the HindIII/XhoI digested vector. To

generate a similar construct which allowed the expression of GFP^{LT} alone, the GFP^{LT} PCR product was digested with HindIII/XhoI and ligated with the HindIII/XhoI digested vector.

Cell cultures

CHO cells were transfected with the vectors containing hybrid cDNA for the C-GFP^{LT} or the GFP^{LT}-C fusion proteins using the calcium phosphate precipitate method in HEPES-buffered saline³⁵. Stable transfectants were selected using 1000 µg Zeocin/ml (Invitrogen) in the growth medium (DMEM with 1000 mg glucose/l, 10 % foetal bovine serum (FBS), 100 µg penicillin-streptomycin mixture ml⁻¹, 2 mM L-glutamine purchased from Life Technologies Inc., Gaithersburg, MD, USA). Untransfected CHO cells were used as the control. To assess the effect of glucagon on fusion protein redistribution, the constructs were stably expressed in BHK/GR cells (Novo Nordisk, Bagsværd, Denmark) overexpressing the human GR. Untransfected BHK/GR cells were used as the control. Expression of GR was maintained with 500 µg G418/ml (*Neo* marker) and C-GFP^{LT} was maintained with 500 µg Zeocin/ml (*Sh ble* marker). CHO cells were also simultaneously co-transfected with vectors containing cDNAs for C-GFP^{LT} and the human ARα2a (ATCC). Transfected cells are referred to as *e.g.* CHO/C-GFP^{LT} cells in the text.

For fluorescence microscopy, cells were allowed to adhere to Lab-Tek chambered coverglasses (Nalge Nunc Int., Naperville, IL, USA) for at least 24 hours and cultured to about 80% confluence. Prior to experiments, the cells were cultured over night without selection pressure in HAM's F12 medium with glutamax (Life Technologies), 100 µg penicillin-streptomycin mixture ml⁻¹ and 0.3 % FBS. This medium has low autofluorescence enabling fluorescence microscopy of cells straight from the incubator.

Immunoblotting

Samples containing 10 µg of protein, determined according to the method of Bradford³⁶ using the Bio-Rad Protein Assay (Bio-Rad Laboratories, Hercules, CA, USA), were added to SDS sample buffer³⁵ and run on precast 7.5 % SDS-PAGE gels with a 4 % stacking gel (Bio-Rad). The proteins were transferred to PH79 nitrocellulose membranes (Schleicher & Schuell GmbH., Dassel, Germany) for an hour at 4°C using a Bio-Rad Transfer Blot apparatus (80 V). Non-specific adhesion was blocked by

incubating the membranes over night in 3 % bovine serum albumin Fraction V (Sigma Chemical Company, St. Louis, MO, USA) in Tris-buffered saline (TBS) containing 50 mM Tris pH 7.5 and 0.15 M NaCl and for an hour in 3 % skim milk powder (Difco Laboratories, Detroit, MI, USA) in TBS with 0.1 % Tween20 (TBST). The membranes were incubated for an hour in TBST with 3 % skim milk powder and the primary polyclonal rabbit anti-C α antibody (Upstate Biotechnology Inc., Lake Placid, NY, USA), which was raised against a peptide corresponding to a 16 amino acid N-terminal stretch of human C α , diluted 1:1000. After 4 washes of 5 min each with TBST, secondary antibody (horse radish peroxidase-conjugated donkey anti-rabbit immunoglobulin from Amersham International plc, Buckinghamshire, UK) diluted 1:5000 in TBS with 3 % skim milk powder was added and incubated for an hour. After 4 washes in TBST and one in TBS, immunoreactivity was detected by enhanced chemiluminescence (ECL) as described by the manufacturer (Amersham) and exposed on Biomax® MR film (Eastman Kodak Company, Rochester, NY, USA). All the steps were performed at room temperature unless otherwise stated.

Time-lapse recording of fusion protein movement.

Cells were cultured in HAM's F12 medium as described above. The chambers were placed on a temperature regulated microscope stage and kept at 37°C. Fluorescence images were captured using an Axiovert 135 inverted light microscope (Carl Zeiss, Oberkochen, Germany) equipped with a Fluor x40, NA 1.3 oil immersion objective (Zeiss) and a cooled (-40°C) CH1 charged coupled device (CCD) camera (Photometrics Ltd., Tucson, AZ, USA). The microscope was equipped with a 470 \pm 20 nm excitation filter, a 505 nm dichroic mirror and a 515 \pm 15 nm emission filter (Delta Lys & Optik, Lyngby, Denmark). The excitation light source was a 100W HBO arc lamp.

Redistribution of the C-GFP^{LT} fusion protein was quantified using an image analysis program custom written in LabVIEW (National Instruments, Austin, TX, USA). Fluorescent aggregates are segmented from each image using an automatically found threshold based on maximisation of the information measure between the object and the background. The *a priori* entropy of the image histogram is used as the information measure³⁷. The area occupied by aggregates in each image is calculated by counting pixels in the segmented areas. The value thus obtained for each image in a series, or treatment pair, is normalised to the value found for the first (unstimulated) image

collected. A value of zero (0) indicates no redistribution of fluorescence from the starting condition. A value of one (1) by this method equals full redistribution.

Spot photobleaching

A Zeiss LSM 410 with x40 Fluar (as above) was used in spot scan mode at 488 nm to bleach individual fluorescent C-GFP^{LT} aggregates within CHO cells variously treated with forskolin. Fluorescence recovery at the locus of each aggregate was monitored immediately after bleach with successive small-area raster scans just large enough to include most of the cell in which the aggregate lay. Nominal output of the laser at 488 nm, before launch into the microscope, was 10 mW. Subsequent raster scans were also run with the laser at full intensity and without a confocal aperture to allow the first to be made within 0.2 seconds of bleach, and for each scan to be completed within 0.3 seconds (100 x 100 pixels per scan). The recovery of fluorescence for the majority of bleach experiments was measured over a period of 215 seconds, recorded in three consecutive blocks of 10 scans having successive intervals between frames of 0.5, 1 and 5 seconds, and a final set of 15 scans each 10 seconds apart. A single scan collected prior to each bleach exposure served both to establish depth of bleach and to estimate maximum recoverable fluorescence in each experiment. Bleach recovery scans (8-bit images) were analysed using IPlab Spectrum software (Signal Analytics Corp., Vienna, VI, USA). A small region of interest (ROI) of between 6x6 to 10x10 pixels was used to define the area for which fluorescence recovery would be monitored in each experiment, and the average fluorescence within that ROI was measured for successive frames in each time series. The measurement ROIs were slightly larger than the bleached C-GFP^{LT} aggregates to allow for cytoplasmic movements during the measurement period. The total average fluorescence within each frame was also measured to allow fluorescence recovery within C-GFP^{LT} aggregates to be corrected for the minor effects of photobleaching caused by the series of measurement scans.

Results of the spot-bleach experiments are presented as normalised values of displacement from photobleach, $\Delta F(t)$, versus time t :

$$\Delta F(t) = [F(\infty) - F(t)]/[F(\infty) - F(0)]$$

where $F(\infty) = F_i R_i / R_i$

$F(\infty)$ being the maximum recoverable fluorescence within a measurement ROI calculated from the pre-bleach intensity of the target aggregate, F_i , corrected for total loss of fluorophores within the cell, R_i/R_o , during the bleach exposure and recovery periods.

SPA

CHO/C-GFP^{LT} cells were cultured in HAM's F12 medium as described above, but in 96-well plates. The medium was exchanged with Ca²⁺-HEPES buffer containing 100 μ M IBMX. The cells were stimulated with different concentrations of forskolin for 10 min. Reactions were stopped with addition of NaOH to 0.14 M and the amount of cAMP produced was measured with the cAMP-SPA kit, RPA538 (Amersham) as described by the manufacturer.

Automated imaging

A Diaphot300 microscope (Nikon Corp., Tokyo, Japan) coupled to a camera based on the SITE back illuminated 512 x 512 CCD camera (Princeton Instruments Inc., Trenton, NJ, USA) and integrated with a digital data acquisition system using LabVIEW software was configured to allow automated focusing and image-based analyses in 96-well plates. CHO/C-GFP^{LT} cells were cultured as described above but in 96-well plates and kept at 37°C throughout the experiments. A fluorescence micrograph of the same field of cells, initially chosen at random, was acquired before and 30 min after forskolin stimulation and analysed as described above.

Endomembrane labelling with fluorescently tagged ceramides

Golgi membranes in CHO/C-GFP^{LT} cells were labelled with ceramide-FL (Molecular Probes Inc., Eugene, OR, USA) at 0.5 μ M for 20 minutes before washing. Ceramide-FL excited at 480 nm normally emits in the green at about 510 nm, but when concentrated (as in Golgi membranes) the fluorophore forms excimers, resulting in a shift in the emission maximum to greater than 600 nm³⁶. Images were collected at both 520 ± 10 nm and beyond 570 nm, allowing good separation of GFP^{LT} and ceramide-FL signals.

Structure of the GFP^{L7}-bright aggregates

Through-focus images of individual C-GFP^{L7} aggregates were collected from chilled cells with a x63 NA 1.4 oil-immersion objective. The built-in focus motor of the Zeiss LSM 410 was used to advance the objective 0.2 μ m between images, 25 images per data set. Effective pixel size in the images was 65.6 nm. Data sets were corrected for bleaching and fluctuations in illumination intensity. Out-of-focus information in the images was removed using iterative, constrained, three-dimensional deconvolution (DeltaVision from Applied Precision Inc., Seattle, WA, USA) based on a theoretically calculated point-spread function. The deconvolved images were then reconstructed into a 3-D rotational projection of 40 images (9 degrees between images) using the method of maximum intensity ray-tracing (DeltaVision, Applied Precision, Inc., Seattle, USA). Two adjacent images in this set, re-sized and pixel-smoothed, were used to create the stereo pair shown in Figure 6C. An iso-surface rendering of the 3-D reconstruction was created using Milan software (BitPlane AG, Zurich, Switzerland) (Fig. 6B).

Acknowledgments

We acknowledge Dr. S. P. Bjørn (Novo Nordisk) for developing a mutagenised cDNA clone for the highly fluorescent derivative of green fluorescent protein used in this study, Dr. G. S. McKnight, Howard Hughes Medical Research Institute, Seattle, USA for providing us with his cDNA clone of the C α subunit of murine cAK and G. Hagel for expert technical assistance.

References

1. Walsh, D.A., Perkins, J.P., and Krebs, E.G. 1968. An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle. *J. Biol. Chem.* **243**:3763-3765.
2. Taskén, K., Skålhegg, B.S., Taskén, K.A., Solberg, R., Knutsen, H.K., Levy, F.O., Sandberg, M., Ørstavik, S., Larsen, T., Johansen, A.K., Vang, T., Schrader, H.P., Reinton, N.T.K., Torgersen, K.M., Hansson, V., and Jahnsen, T. 1997. Structure, function, and regulation of human cAMP-dependent kinases. *Adv. Second Messenger Phosphoprotein Res.* **31**:191-204.
3. Kingston, P.A., Zufall, F., and Barnstable, C.J. 1996. Rat hippocampal neurons express genes for both rod retinal and olfactory cyclic nucleotide-gated channels: Novel targets for cAMP/cGMP function. *Proc. Natl. Acad. Sci. USA* **93**:10440-10445.
4. Taskén, K., Solberg, R., Foss, K.B., Skålhegg, B.S., Hansson, V., and Jahnsen, T. 1995. *The Protein Kinase Facts Book: Protein-Serine Kinases* (Eds., Hardie, G., and Hanks, S.), Academic Press Ltd., London, UK, pp 58-63
5. Taskén, K., Skålhegg, B.S., Solberg, R., Andersson, K.B., Taylor, S.S., Lea, T., Blomhoff, H.K., Jahnsen, T., and Hansson, V. 1993. Novel isozymes of cAMP-dependent protein kinase exist in human cells due to formation of RI α -RI β heterodimeric complexes. *J. Biol. Chem.* **268**:21276-21283.
6. Beebe, S.J., Salomonsky, P., Jahnsen, T., and Li, Y. 1992. The Cy subunit is a unique isozyme of the cAMP-dependent protein kinase. *J. Biol. Chem.* **267**:25505-25512.
7. Gamm, D.M., Baude, E.J., and Uhler, M.D. 1996. The major catalytic subunit isoforms of cAMP-dependent protein kinase have distinct biochemical properties *in vitro* and *in vivo*. *J. Biol. Chem.* **271**:15736-15742
8. Robinson, M.L., Wallert, M.A., Reinitz, C.A., and Shabb, J.B. 1996. Association of the type I regulatory subunit of cAMP-dependent protein kinase with cardiac myocyte sarcolemma. *Arch. Biochem. Biophys.* **330**:181-187.
9. Moos, J., Peknicová, J., Geussová, G., Philimonenko, V., and Hozák, P. 1998. Association of protein kinase A type I with detergent-resistant structures of mammalian sperm cells. *Mol. Reprod. Dev.* **50**:79-85.

10. Nigg, E.A., Schäfer, G., Hilz, H., and Eppenberger, H.M. 1985. Cyclic-AMP-dependent protein kinase type II is associated with the Golgi complex and with centrosomes. *Cell* **41**:1039-1051.
11. Li, Y., Ndubuka, C., and Rubin, C.S. 1996. A kinase Anchor protein 75 targets regulatory (RII) subunits of cAMP-dependent protein kinase II to the cortical actin cytoskeleton in non-neuronal cells. *J. Biol. Chem.* **271**:16862-16869.
12. Pariset, C., and Weinman, S. 1994. Differential localization of two isoforms of the regulatory subunit RII α of cAMP-dependent protein kinase in human sperm: Biochemical and cytochemical study. *Mol. Reprod. Dev.* **39**:415-422.
13. Shmyrev, I.I., Grozdova, I.D., Kondratyev, A.D., Mamayeva, E.G., and Severin, E.S. 1990. Immunofluorescence localization of the regulatory subunit type II of cAMP-dependent protein kinase in PC12 and 3T3 cells in different proliferative states. *Mol. Cell. Biochem.* **93**:47-52.
14. Zhang, Q., Carr, D.W., Lerea, K.M., Scott, J.D., and Newman, S.A. 1996. Nuclear localization of type II cAMP-dependent protein kinase during limb cartilage differentiation is associated with a novel developmentally regulated A-kinase anchoring protein. *Dev. Biol.* **176**:51-61.
15. Coghlan, V.M., Bergeson, S.E., Langeberg, L., Nilaver, G., and Scott, J.D. 1993. A-Kinase Anchoring Proteins: a key to selective activation of cAMP-responsive events? *Mol. Cell. Biochem.* **127/128**:309-319.
16. Huang, L.J.-S., Durick, K., Weiner, J.A., Chun, J., and Taylor, S.S. 1997. D-AKAP2, a novel protein kinase A anchoring protein with a putative RGS domain. *Proc. Natl. Acad. Sci. USA* **94**:11184-11189.
17. Carr, S.A., Biemann, K., Shoji, S., Parmelee, D.C., and Titani, K. 1982. *n*-Tetradecanoyl is the NH₂-terminal blocking group of the catalytic subunit of cyclic AMP-dependent protein kinase from bovine cardiac muscle. *Proc. Natl. Acad. Sci. USA* **79**:6128-6131.
18. Adie, E.J., Thomas, P.H., Munday, M.R., and Clegg, R.A. 1995. Subcellular targeting of recombinant and mammalian C α subunits of cAMP-dependent protein kinase. *Biochem. Soc. Trans.* **23**:451S.
19. Yang, S., Fletcher, W.H., and Johnson, D.A. 1995. Regulation of cAMP-dependent protein kinase: Enzyme activation without dissociation. *Biochemistry* **34**:6267-6271.

20. Hardie, D.G. 1991. *Biochemical Messengers: Hormones, Neurotransmitters and Growth Factors*, Chapman & Hall, London, UK
21. Harootunian, A.T., Adams, S.R., Wen, W., Meinkoth, J.L., Taylor, S.S., and Tsien, R.Y. 1993. Movement of the free catalytic subunit of cAMP-dependent protein kinase into and out of the nucleus can be explained by diffusion. *Mol. Biol. Cell* **4**:993-1002.
22. Nigg, E.A., Hilz, H., Eppenberger, H.M., and Dutly, F., 1985. Rapid and reversible translocation of the catalytic subunit of cAMP-dependent protein kinase type II from the Golgi complex to the nucleus. *EMBO J.* **4**:2801-2806.
23. Walsh, D.A., and Van Patten, S.M. 1994. Multiple pathway signal transduction by the cAMP-dependent protein kinase. *FASEB J.* **8**:1227-1236.
24. Mellon, P.L., Clegg, C.H., Correll, L.A., and McKnight, G.S. 1989. Regulation of transcription by cyclic AMP-dependent protein kinase. *Proc. Natl. Acad. Sci. USA* **86**:4887-4891.
25. Hagiwara, M., Brindle, P., Harootunian, A., Armstrong, R., Rivier, J., Vale, W., Tsien, R., and Montminy, M.R. 1993. Coupling of hormonal stimulation and transcription via the cyclic AMP-responsive factor CREB is rate limited by nuclear entry of protein kinase A. *Mol. Cell. Biol.* **13**:4852-4859.
26. Adams, S.R., Harootunian, A.T., Buechler, Y.J., Taylor, S.S., and Tsien, R.Y. 1991. Fluorescence ratio imaging of cyclic AMP in single cells. *Nature* **349**:694-697.
27. Elson, E.L., and Qian, H. 1989. Interpretation of fluorescence correlation spectroscopy and photobleaching recovery in terms of molecular interactions. *Methods Cell Biol.* **30**:307-332.
28. Lipsky, N.G., and Pagano, R.E. 1985. A vital stain for the Golgi apparatus. *Science* **288**:745-747.
29. Knighton, D.R., Zheng, J., Ten Eyck, L.F., Ashford, V.A., Xuong, N.-h., Taylor, S.S., and Sowadski, J.M. 1991. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**:407-414.
30. Knighton, D.R., Zheng, J., Ten Eyck, L.F., Xuong, N.-h., Taylor, S.S., and Sowadski, J.M. 1991. Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253**:414-420.

31. Scott, J.D. 1991. Cyclic nucleotide-dependent protein kinases. *Pharmacol. Ther.* **50**:123-145.
32. Skålhegg, B.S., Keryer, G., Torgersen, K.M., Aandahl, E.M., Levy, F.O., Jahnsen, T., Hansson, V., and Taskén, K. 1997. Function and localization of cAMP-dependent protein kinase type I and II in lymphoid cells. *FEBS special meeting 1997: Cell signalling mechanisms*, Elsevier Science, The Netherlands, Abstract P6-123
33. Akey, C.W. 1995. Structural plasticity of the nuclear pore complex. *J. Mol. Biol.* **248**:273-293.
34. Uhler, M.D., Carmichael, D.F., Lee, D.C., Chrivia, J.C., Krebs, E.G., and McKnight, G.S. 1986. Isolation of cDNA clones coding for the catalytic subunit of mouse cAMP-dependent protein kinase. *Proc. Natl. Acad. Sci. USA* **83**:1300-1304.
35. Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual, 2nd Ed.*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA
36. Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-254.
37. Kapur, J.N., Sahoo, P.K., and Wong, A.K.C. 1985. A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing* **29**:273-285.
38. Haugland, R.P. 1996. *Handbook of fluorescent probes and research chemicals*, 6th Ed. Molecular Probes, Oregon, USA, p281.

Figure legends

Table 1. Time from initiation of a response to half maximal ($t_{1/2max}$) and maximal (t_{max}) C-GFP^{LT} redistribution. The data was extracted from curves such as shown in Figure 3B. All $t_{1/2max}$ and t_{max} values are given as mean \pm SD and are based on a total of 26-30 cells from 2-3 independent experiments for each forskolin concentration. Since the observed redistribution is sustained over time, the t_{max} values were taken as the earliest time point at which complete redistribution is reached. Note that the values do not relate to the degree of redistribution.

Figure 1. Western blot analysis of lysates containing C-GFP^{LT} fusion proteins. Total lysates of BHK/GR,C-GFP^{LT} (A) and control BHK/GR (B) cells were probed with an anti-C α antibody. 500 ng of purified bovine C subunit (C) was included as a positive control and to identify the endogenous C subunit. Although the antibody clearly reacts unspecifically with several proteins in the total lysates, the fusion protein (f) is recognised as a specific band, migrating with an apparent size of 60 kDa, in the transfected cells (A). The endogenous C subunit (e) migrated as predicted by its molecular weight of 41 kDa. It is possible to compare the expression levels of endogenous hamster C subunit and overexpressed mouse fusion proteins in these blots since the immunogenic peptide is conserved between these two species.

Figure 2. Fluorescence micrographs of CHO cells expressing C subunit fusion proteins. The two fusion proteins of the C subunit of cAK show distinct localisation patterns. A. The NH₂-terminal GFP^{LT}-C fusion protein is localised almost evenly throughout the cytoplasm. B. The COOH-terminal C-GFP^{LT} fusion protein is highly concentrated in cytoplasmic aggregates, often in one large and several minor structures per cell. Scale bar 10 μ m.

Figure 3. Time-lapse analyses of fluorescence redistribution in CHO/C-GFP^{LT} cells treated with various agonists. The raw data of each experiment consisted of 60 fluorescence micrographs acquired at regular intervals including several images acquired before the addition of agonist. Six of these images are shown (A) for the typical response to 1 μ M forskolin, taken at the time points indicated. The time point t=0 corresponds to the image acquired immediately before the cells were challenged with agonist. Scale bar

10 μm . The charts (B-G) each show a quantification of the responses in each time series. The total area of the highly fluorescent aggregates (see Experimental Protocol) is plotted versus time for each experiment. (B) Redistribution time profiles of the C-GFP^{LT} fusion following treatment of cells with various concentrations of forskolin. (C) Response following addition of 1 mM dbcAMP. (D) The effect of 100 μM IBMX on the fusion protein distribution. (E) Demonstrates the reversibility of the forskolin-induced redistribution of C-GFP^{LT}, where 10 μM forskolin (open arrow) is followed shortly by repeated washings with buffer (dark arrow). In a parallel experiment, treatment with 10 μM forskolin plus 100 μM IBMX is followed by repeated washing with buffer containing 100 μM IBMX. (F) BHK/GR,C-GFP^{LT} cells treated with 100 nM glucagon. (G) CHO/C-GFP^{LT} cells transiently transfected with the AR α 2a were pretreated with 10 μM forskolin (open arrow) to increase [cAMP], then given 10 μM norepinephrine in the continued presence of forskolin.

Figure 4. (A) Four frames from the recovery sequence following spot photobleach of a large aggregate (arrow) in a CHO/C-GFP^{LT} cell exposed to 25 μM forskolin. Times are seconds after bleach. (B) Normalised displacement curves of the fluorescence recovery process in cells exposed to various levels of forskolin. Measurement points are averages \pm sem (n=4). (C) Linear fits to the first five points of the normalised recovery curves shown in (B). The slope of each line is used as an estimate of the half-time of recovery from bleach at each forskolin concentration.

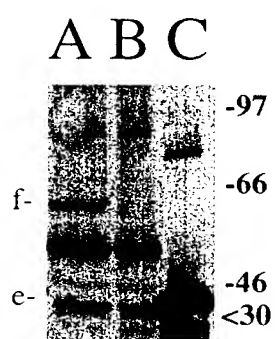
Figure 5. Parallel dose response analyses of forskolin effects in CHO/C-GFP^{LT} cells on: [cAMP]_i elevation (\square), the rate of recovery from spot photobleach (Δ) and induced change in C-GFP^{LT} redistribution (\bullet). [cAMP]_i was measured by SPA assay, analysing the effects of buffer or 8 increasing concentrations of forskolin in these cells. The graph shows a trace of the mean \pm sem expressed in arbitrary units (n=4 for each data point). Half times for recovery from spot photobleach were estimated from the first 5 time points of the mean value (n=4) curves in Figure 4B. Changes induced in C-GFP^{LT} distribution were quantified as described (Experimental Protocol) using fluorescence micrographs taken of the same field of cells prior to and 30 min after the addition of forskolin. The graph shows a trace of the mean \pm sem at each forskolin concentration (n=8 for each data point). The fitted curves indicate $\frac{1}{2}$ -maximal concentration values for

forskolin as: 1.7 μ M, image-based assay (\square); 3.0 μ M, spot photobleach assay (Δ); 9.3 μ M, SPA (\bullet).

Figure 6. (A) Two images of CHO/C-GFP^{LT} cells stained with ceramide-FL, in emission ranges of 520 ± 10 nm and >570 nm, have been superimposed to demonstrate the distinct separateness of Golgi membranes (orange) and C-GFP^{LT} fluorescence (green). Scale bar is 10 μ m. (B) An iso-surface rendering of a single large C-GFP^{LT} aggregate (similar to that arrowed in 6A). The image is a reconstruction from 25 through-focus images deconvolved and processed as described (Experimental Protocol). Scale bar 1 μ m. (C) Stereo pair of the reconstructed images used to generate the iso-surface seen in (B). Each image is smoothed for presentation, the structure originally being 35 pixels high by 27 wide in this orientation. Scale bar 1 μ m.

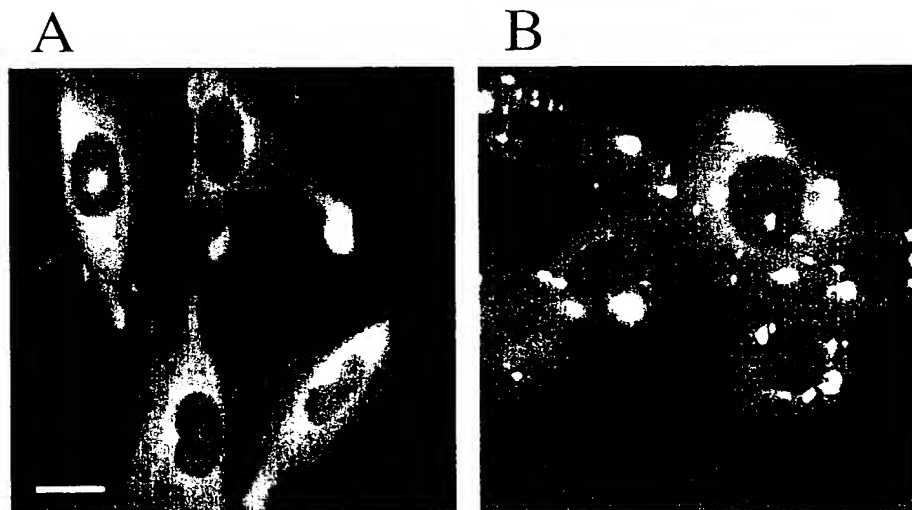
Modtaget PD
15 OKT. 1998

Figure 1



Modtaget PD
15 OKT, 1998

Figure 2



Modtaget PD
15 OKT. 1998

Figure 3

A

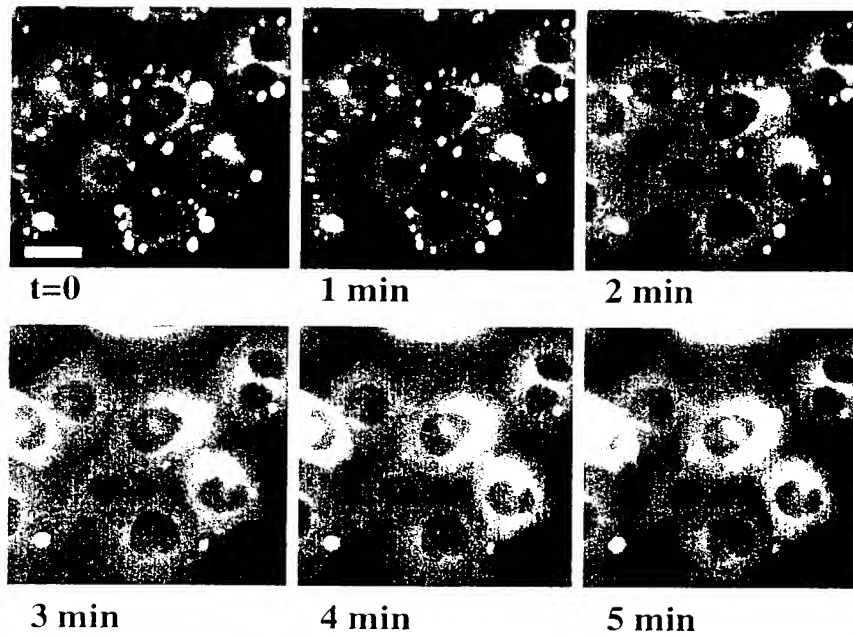


Figure 3

Modtaget PC
15 OCT. 1998

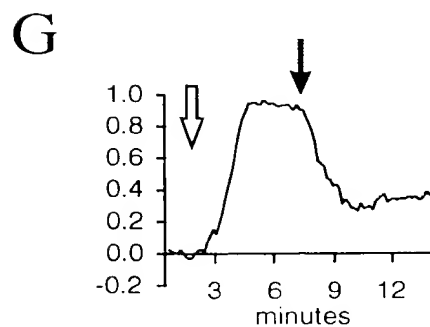
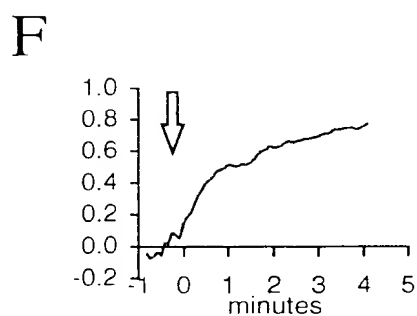
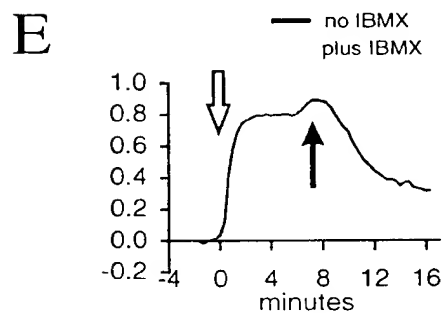
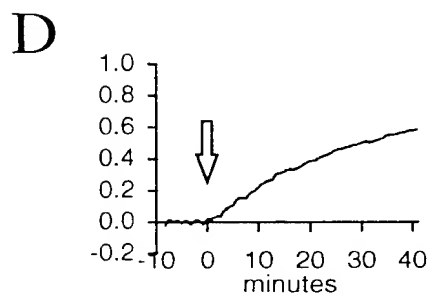
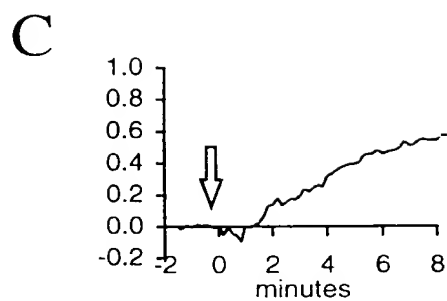
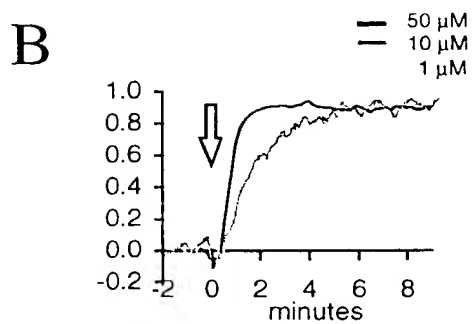
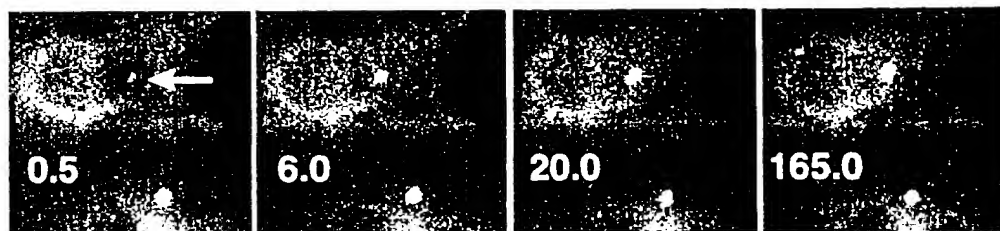
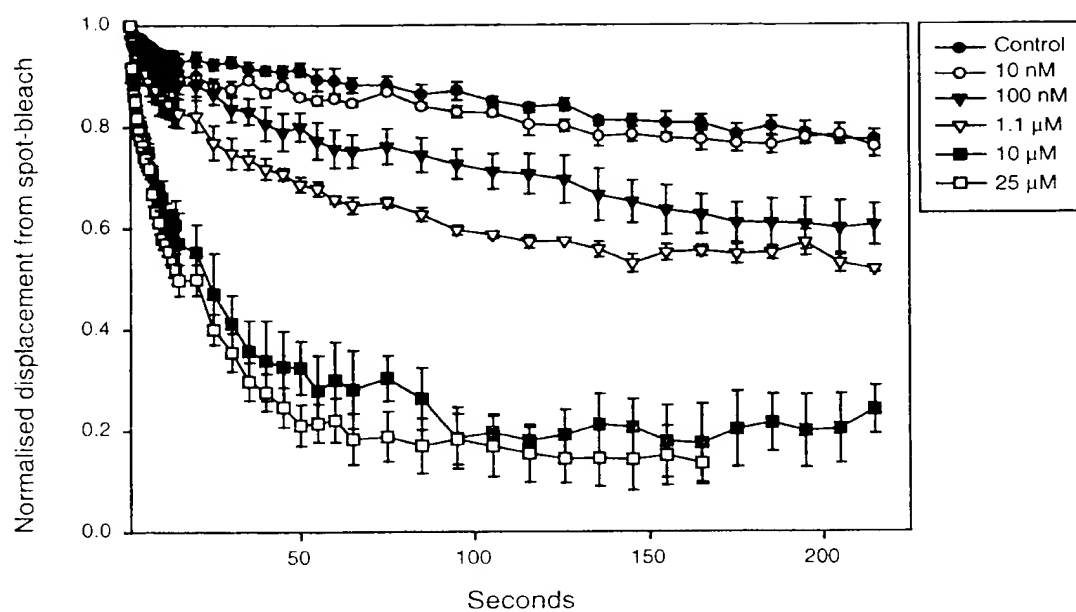


Figure 4

A



B



C

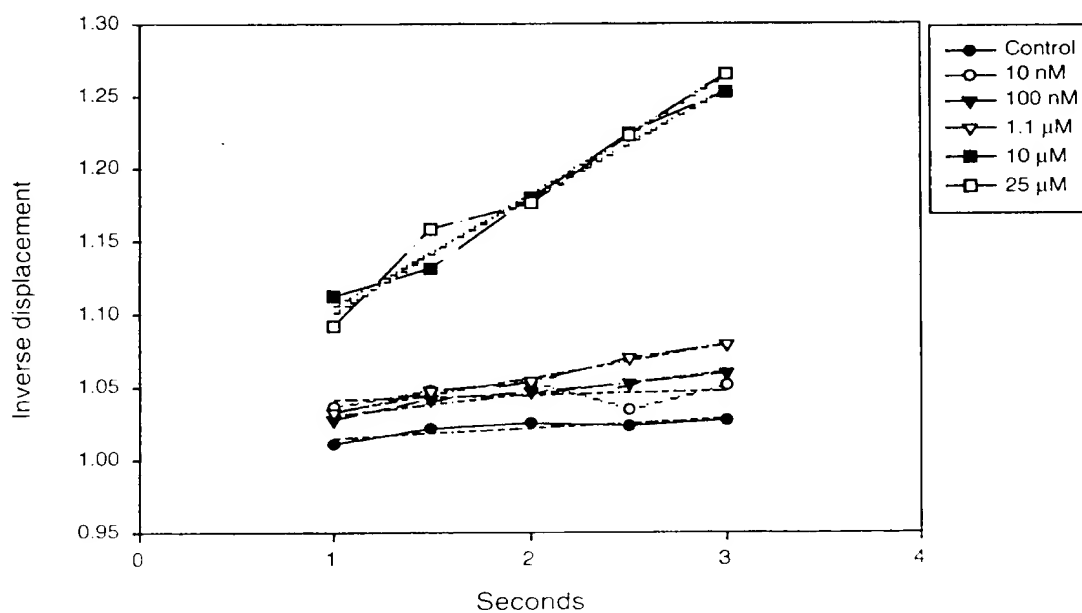
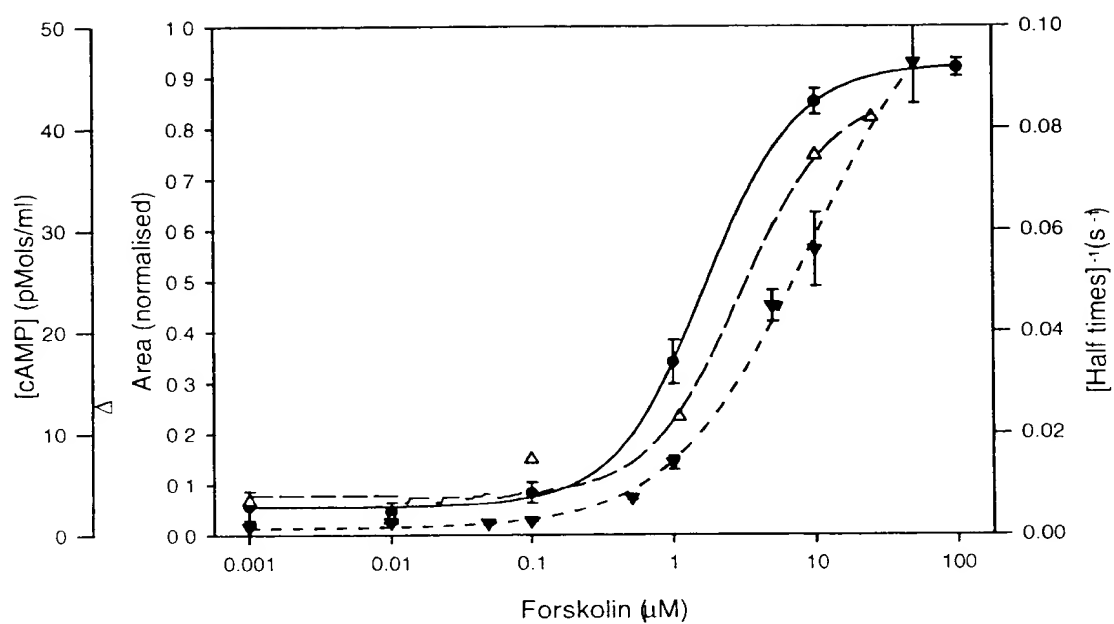


Figure 5

Modtag
18 5 OK1 19...



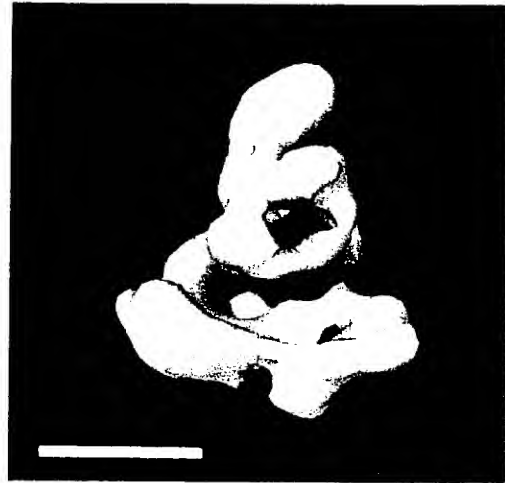
Modtageri PE
35 OKT. 1998

Figure 6

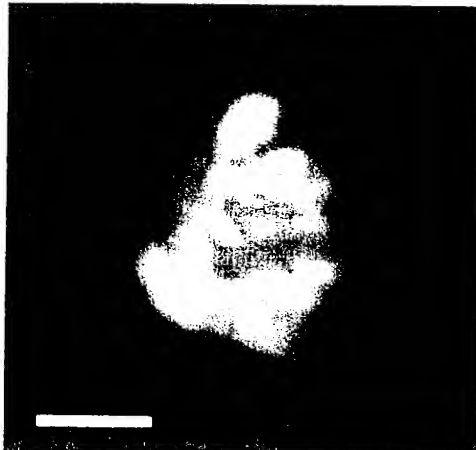
A



B



C



Left

Right

Modtaget PD
15 OKT. 1998

Table 1

[forskolin]/ μ M	$t_{1/2\max}$ / s	t_{\max} / s
1	115 \pm 21	310 \pm 31
10	69 \pm 14	224 \pm 47
50	47 \pm 10	125 \pm 28